DOCUMENT RESUME

ED 161 706 SE 025 160

TITLE	Minicourses in Astrophysics, Modular Approach, Vol. II.
INSTITUTION	Illinois Univ., Chicago.
SPONS AGENCY	National Science Foundation, Washington, D.C.
BUREAU NO	SED-75-21297
PUB DATE	77
NOTE	138p.; For related document, see SE 025 159; Contains occasional blurred, dark print
EDRS PRICE	MF-$0.83 HC-$7.35 Plus Postage.
DESCRIPTORS	*Astronomy; *Curriculum Guides; Evolution; Graduate Study; *Higher Education; *Instructional Materials; Light; Mathematics; Nuclear Physics; *Physics; Radiation; Relativity; Science Education; *Short Courses; Space Sciences
IDENTIFIERS	*Astrophysics

ABSTRACT

This is the second of a two-volume minicourse in astrophysics. It contains chapters on the following topics: stellar nuclear energy sources and nucleosynthesis; stellar evolution; stellar structure and its determination; and pulsars. Each chapter gives much technical discussion, mathematical treatment, diagrams, and examples. References are included with each chapter. (BB)

MINICOURSES IN ASTROPHYSICS
MODULAR APPROACH
VOL. II

DEVELOPED AT THE
UNIVERSITY OF ILLINOIS AT CHICAGO
1977

SUPPORTED BY NATIONAL SCIENCE FOUNDATION

DIRECTOR: S. SUNDARAM
DEPARTMENT OF PHYSICS
UNIVERSITY OF ILLINOIS
CHICAGO, ILLINOIS 60680

ASSOCIATE DIRECTOR: J. BURNS
DEPARTMENT OF PHYSICS AND SPACE SCIENCES
FLORIDA INSTITUTE OF TECHNOLOGY
MELBOURNE, FLORIDA 32901

STELLAR NUCLEAR ENERGY SOURCES

and

NUCLEOSYNTHESIS

I. Introduction

In this minitext the subject of nuclear reactions in stars will be surveyed with three aims in mind: to show what nuclear processes play important roles in generating the energy radiated by stars during the principal stages of their evolution, to see what physical factors govern the rates of energy production by these processes, and finally to demonstrate the roles these processes play in building up heavy elements from the main primordial material of the Universe, hydrogen and helium.

Two quite different sources of energy supply the energy radiated by a star: gravity and nuclear fusion. As a rule these sources alternate in importance in successive steps in the evolution of a star from birth to its final state. Initially gravitational contraction condenses and heats interstellar gas and dust to form a star, and when this process has raised interior temperatures into the range of $\approx 10^7$°K, fusion of hydrogen into helium by various processes takes over the job of providing the star's energy. At this point gravitational contraction ceases and gravity no longer plays an active role in energy generation. However, when the hydrogen in the central region of the star has been converted into helium, the fuel for the initial forms of nuclear energy generation is exhausted. These fusion processes then cease and gravity takes over, causing the star to contract and heat further until temperatures of the order of 10^8°K are reached. At these temperatures helium nuclei can fuse into carbon nuclei with release of energy, and a new source of nuclear energy enters the picture.

Eventually the fuel for this process is exhausted. Gravity steps in once more and heats the star until another exothermic nuclear reaction becomes possible, and the process repeats in this fashion until it is interrupted by one of several possible terminal steps which will be discusse later.

The key factor in production of energy by nuclear fusion is the fact that heavier nuclei are bound more tightly than lighter nuclei, and when they form from fusion of lighter nuclei, the difference in binding energie of the initial and final nuclei is given off in the form of energetic electrons, positrons, gamma rays, and neutrinos. The common manifestation of this factor can be found in the familiar chemical periodic table of the elements where the mass of a helium atom is less than the total mass of four hydrogen atoms that can fuse to form one helium. This mass defect, a it is called, is evidence of the binding energy increase, and the energy given off in the fusion process is $E = \Delta mc^2$ by the Einstein relation.

To give an explicit example, consider the fusion reaction

$$4H^1 \rightarrow He^4$$

$4 \times 1.008 \rightarrow$	4.003 +	$.0029$ gm
mass of 4 moles of hydrogen	mass of 1 mole of helium	mass equivalent of the excess energy

The excess energy equivalent .0029 gm/mole of He formed is equal by the Einstein relation to 2.61×10^{18} ergs, or 6.5×10^{17} ergs/gm of He formed. This is quite a lot of energy; in fact it is approximately 10^5 times the energy released in a typical chemical reaction, even so energetic a reaction as the explosive burning of hydrogen in air. Such an energetic proce is necessary, however, to supply the observed energy production rates of stars.

The Sun, for example, radiates 4×10^{33} ergs/sec from a mass of 2×10^{33} gms., or about 2 ergs/sec per gram. It can continue at this rate for some 10 billion years or $\approx 3 \times 10^{17}$ sec, yielding a total of 6×10^{17} ergs/gm, in good agreement with the value calculated above for energy production from hydrogen fusion.

Gravitational contraction is an efficient energy source, but it operates on a much shorter time scale than nuclear energy production. For example, gravitational energy released by contraction of the Sun from an initial gas cloud to its present size would fuel the Sun at its present rate of radiation for only 20 million years. But geological evidence on Earth shows that the Sun has been radiating at its present rate for at least 3.5 billion years, and present theory of stellar evolution predicts that a star like the Sun will radiate at nearly its present rate for approximately 10 billion years before undergoing any substantial changes. Therefore, nuclear burning stages in the life of a star tend to be relatively long and stable while gravitational energy production stages are relatively short and often dynamically unstable.

The general processes of fusion lead to production of heavier elements from lighter ones. The building up of heavier elements proceeds roughly in the sequence $H \rightarrow He \rightarrow C \rightarrow N \rightarrow O \rightarrow F \rightarrow Ne \rightarrow Mg \rightarrow Si$ ..... Fe. Many steps and details have been omitted in this outline, and some of these will be filled in later. There are several features of this bare outline of nucleosynthesis that deserve comment at this point. First, the sequence stops at Fe^{56}. This nucleus is the most stable one in nature, and the construction of any heavier nucleus than Fe^{56} requires a net input of energy rather than giving off energy, so energy production by fusion stops at Fe^{56}. It is possible to make heavier elements than iron in stars but the process becomes more difficult and requires rather special conditions, as we shall see.

The second point to notice in the building up scheme above is that there are lots of gaps in the periodic table of elements made by simple fusion. For instance, there is no Na^{23} or Be^{8} or Li^{7}, etc., in the sequence above. These elements are actually produced in small amounts in some of the nuclear reactions that give the main products listed above, and we shall show how these minor constituents are created. Toward the heavier end of the sequence there are a number of nuclei that are somewhat neutron-rich in that they have more neutrons than protons. These elements cannot be created without the presence of free neutrons, and some of the steps in production of intermediate elements just beyond atomic weight ≈20 do occasionally yield neutrons which can be captured to give the isotopes missing from the preceding list.

II. Nuclear Reaction Rates

It is clear that while stars last a long time by our reckoning, they are nevertheless dynamic, evolving objects. Consequently, the rates at which various processes that go on in a star are of primary importance in understanding its behavior. In particular, we will want to know the rates of possible nuclear reactions both from a standpoint of the stars' energy production and in order to understand how elements are synthesized by nuclear reactions in stars.

The rate of a particular nuclear reaction is the product of several factors that are not difficult to understand in broad terms, although some of these factors may be quite complex in detail. In its simplest form, a nuclear reaction may be considered as the possible result of a collision between two particles. These particles may be of various types including nucleons (i.e., protons, helium nuclei (alpha particles) and other nuclei such as deuterium, He^3, C^{12}, etc.), and lighter particles such as electrons, positrons, neutrinos, and gamma ray photons.

It is easy to see that the reaction rate (number of nuclear reactions per second per unit volume) must be proportional to the collision rate, the number of collisions per second per cm^3. This may be calculated in a straightforward way. Let the effective "size" of a particle of one kind, say Type 1, be described by radius r_1 so it may be considered to collide with any other particle it approaches closer than r_1. If this particle moves at velocity v_1, in one second it sweeps out a cylinder in space of height v_1 and base area πr_1^2 which we designate σ_1. This "collision cylinder" of area σ_1, length v_1, is such that the Type 1 particle will collide with anything within it. The area σ_1 is called the collision cross section of a Type 1 particle. The cylinder swept out in one second has a volume $\sigma_1 v_1$. (See Figure 1)

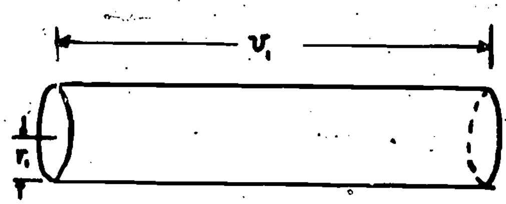


Fig. 1. Collision cylinder swept out in one second by a Type 1 particle.

If there are N_2 particles of Type 2 per cm^3 in the gas, a Type 1 particle will collide with $\sigma_1 v_1 N_2$ Type 2 particles each second, and if there are N_1 particles of Type 1 per cm^3 the total collision rate will be $\sigma_1 v_1 N_1 N_2$. This example was based on the assumption that every Type 1 particle moved at velocity v_1 and that all Type 2 particles are at rest. In reality both types of particles have a wide range of velocities, and a suitable average of the product (σv) must be made over the distribution of particle velocities characterized by the gas temperature T. This velocity distribution is the Maxwell-Boltzman distribution which gives the fraction of particles, n(v), having speeds between v and v + dv

$$n(v)\,dv = 4\pi \left(\frac{m}{2\pi kt}\right)^{3/2} v^2 e^{-mv^2/2kt}\,dv \qquad (1)$$

Collisions between nuclear particles take place on so small a size scale ($\approx 10^{-13}$ cm) that quantum effects are very important, and the simple classical picture above must be modified. The principal change is that nuclear dimensions such as r_1 must be expressed in terms of the corresponding deBroglie wavelength, h/mv. This means that the collision cross section σ will be proportional to the square of the deBroglie wavelength, i.e., to v^{-2}, and the product σv in the collision rate must be replaced by v^{-1}. This makes the collision rate R_{12} proportional to v^{-1}, or

$$R_{12} = AN_1N_2\,v^{-1} \text{ collisions per cm}^3 \text{ per second} \qquad (2)$$

where A is a constant that includes Planck's constant and the masses of the colliding particles. It would be premature to average (2) over the velocity distribution (1) at this point because there are more velocity dependent factors yet to be included in the overall reaction rate.

We must turn next to some basic facts of nuclear physics to see how close nuclei must approach each other for a nuclear reaction to take place. The central feature of nuclear forces, for purposes of this discussion, is that they are very short range forces and quite strong. We are accustomed to forces like gravity and electrostatic forces whose potentials fall off slowly with distance in r^{-1} fashion and to dipole forces that fall off as r^{-3} or even vander Waals forces between molecules that go as r^{-7}. The forces between nuclei fall off much more rapidly than any of these, and it is a fairly good approximation to think of the nuclear force as being derived from a "square-well" potential like that shown in Figure 2.

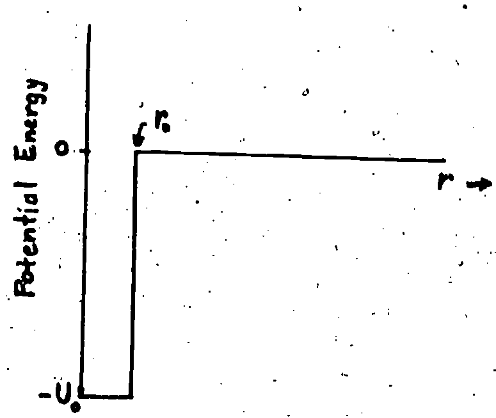


Fig. 2. Nuclear "square-well" potential

The range, r_0, of this nuclear force is of the order of 10^{-13} cm, the size of a nucleon, and the force is attractive, meaning that once two nuclei touch they coalesce much as do two small droplets of liquid that combine under the influence of surface tension forces once their surfaces come into contact. This analogy with a liquid is widely used in nuclear physics, and it is a useful way to think about collisions leading to nuclear reactions.

Most of the nuclear processes of interest in stellar energy production involve interactions between charged nuclei such as protons, alpha particles (doubly charged He nuclei), and so forth. These particles, being positively charged, repel each other strongly at close range, and since they must approach to distances $\approx 10^{-13}$ cm to react, they must have high initial velocities (i.e., high temperatures) to overcome the strong electrostatic Coulomb repulsion. The Coulomb repulsion of two protons separated by 2×10^{-13} cm is $\approx .6$ Mev (1 Mev $= 1.6 \times 10^{-6}$ ergs). This corresponds to two protons colliding head-on each with initial velocity $\approx 1.2 \times 10^{9}$ cm/sec which is .04 times the speed of light. In terms of the average energy, $\frac{3}{2}kT$, of a particle in a gas at temperature T (k is the Boltzmann constant 1.38×10^{-16} ergs/degree), the Coulomb repulsion corresponds with the average energy of hydrogen gas at 5.5×10^{9} degrees K. Such temperatures are far above those found in the interiors of normal stars whose central temperatures range mostly around 10-30 million °K. It is true that a few very high speed particles are present in a gas at a given temperature, even particles whose kinetic energy are as much as 400 times the average particle energy as required for the proton-proton collision in the example. But the fraction of such energetic particles is very small indeed ($\approx e^{-400} \approx 10^{-160}$), much too small to allow nuclear reactions to proceed at a rate sufficient to account for the energy output of a typical star. There must be some way by which the Coulomb barrier can be foiled, and the method was first discovered by Gamow in connection with a related problem, natural radioactive decay of heavy elements by alpha particle emission. Gamow showed that charged particles which do not have enough energy to go <u>over</u> the barrier may instead go <u>through</u> it by the process of quantum mechanical tunnelling as shown schematically in Figure 3 (next page).

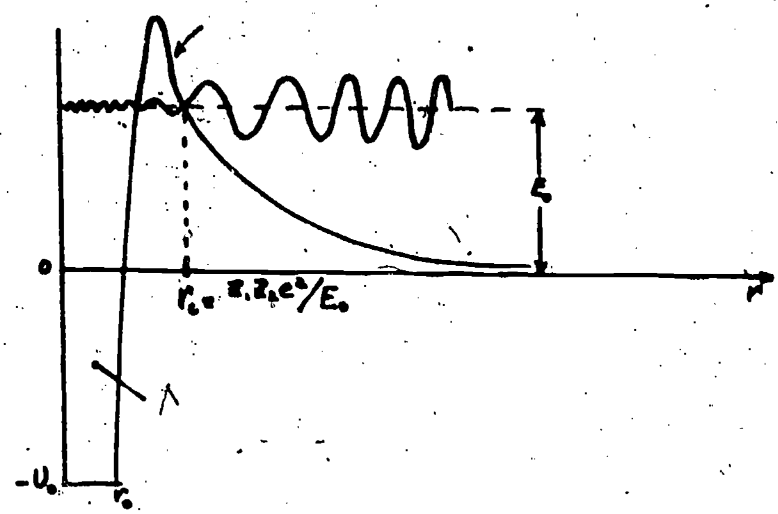


Fig.3. Schematic representation of tunneling of a charged particle through a Coulomb barrier.

Quantum mechanics describes the position of the incident particle in the form of deBroglie waves of wavelength $\propto (E - U(r))^{-1/2}$ which becomes imaginary within the barrier where $E - U(r)$ is negative. Elementary theory of wave propagation shows that such a wave will propagate with exponentially decreasing amplitude into a region where its wavelength is imaginary. This is true of electromagnetic waves (e.g., penetration into metals to the "skin depth") and acoustical waves (penetration into dissipative medium), to cite two familiar examples, and it is also true of matter waves (deBroglie waves). Therefore, the matter waves associated with an incident charged particle penetrate into the barrier between r_o and r_c where the wavelength λ is imaginary, but the amplitude decreases exponentially with distance into the barrier rather than going to zero abruptly at $E = U(r_c)$ where classically the incident particle would stop. This means that there is a small but finite probability for the incident particle to tunnel through the barrier and be found inside r_o. This probability for tunnelling can be calculated quantum mechanically and the result is that the barrier transmission (fraction of incident particles that get through) is

$$T \propto E^{-1/2} \exp\left(-4\pi^2 Z_1 Z_2 e^2/hv\right) \qquad (3)$$

Once the incident particle has penetrated the barrier, there is no certainty that a nuclear reaction will take place: the particles might separate again almost at once, or some other nuclear reaction than the one of interest might take place. A factor, Γ, is therefore included in the reaction rate expression. It is called the "nuclear factor" and is the probability that the desired reaction will occur once the Coulomb barrier has been penetrated. This factor is very difficult to calculate with any accuracy at the present state of knowledge about low energy nuclear reactions, and it is usually determined experimentally.

Putting all the factors together, the overall reaction rate can be expressed in the form

$$R = \int_0^\infty \left(C\, N_1 N_2\, v^{-1}\right) \times \left(v^2 T^{-3/2}\, e^{-\frac{mv^2}{2kT}}\right) \times \left(e^{-\frac{4\pi^2 Z_1 Z_2}{hv}}\right) \Gamma \; dv \qquad (4)$$

Collision factor | Maxwell-Boltzmann velocity distribution | Gamow factor (barrier transmission) | Nuclear factor

It is evident that integration of the above expression gives a somewhat complicated result. For particle energies where Γ does not vary rapidly with energy (i.e., where there are no "resonances" or energy regions of enhanced nuclear reaction probability) the rate of energy production, ε, which is proportional to (4) takes the form

$$\varepsilon = B\rho X_1 X_2\, T^{-2/3}\, e^{-b/T^{1/3}} \qquad (5)$$

where B and b are constants, ρ is the gas density, X_1 and X_2 are the fractions by weight of the interacting nuclear species (Types 1 and 2 in the earlier discussion), and T is the absolute temperature. Even this expression is inconvenient for many uses and can be expressed in still simpler form over a limited temperature range,

$$\varepsilon = A\rho X_1 X_2 T^n \qquad (6)$$

where n is a positive number that varies somewhat with energy but which can be taken as constant over a reasonable temperature range. The exponent n is large ranging from about 4 for the pp cycle described below at temperatures around $2 \times 10^{7\,\circ}K$ to about 20 for the CNO cycle at temperatures around $1.5 \times 10^{7\,\circ}K$.

III. Nuclear Energy Production in Stars

The simplest and most efficient energy producing nuclear reactions involve the building up of He^4 nuclei from H^1 nuclei (protons) by two principle modes. One of these, the p-p chain, involves only light nuclei up to He^4 in its principal form while the other involves carbon, nitrogen, and oxygen nuclei and is commonly referred to as the CN chain or CNO bi-cycle since it consists of two interrelated reaction cycles.

The p-p chain in its principal form (p-p I) consists of three steps that take place in series as shown below:

$$\left.\begin{array}{ll} (1) & H^1 + H^1 \rightarrow D^2 + e^+ + \nu \\ (2) & H^1 + D^2 \rightarrow He^3 + \gamma \\ (3) & He^3 + He^3 \rightarrow He^4 + 2H^1 \end{array}\right\} \text{p-p I chain} \qquad (7)$$

In these reactions D^2 is a deuterium nucleus, e^+ is a positron, γ is a gamma ray, and ν is a neutrino. Other possible reactions such as $H^1 + He^3$,

$D^2 + D^2$, $D + He^3$ proceed at negligible rates and have been omitted. It is evident that the overall effect of this reaction chain is to fuse four protons into an alpha particle:

$$4\,H^1 \rightarrow He^4 + 2e^+ + 2\nu + 2\gamma \qquad (8)$$

The 2 positrons soon find electrons, annihilate, and yield 2 more γ's each for a total 6 gamma rays plus 2 neutrinos emitted for each alpha particle (He^4) formed.

There are two additional related groups of reactions that temporarily form nuclei of somewhat higher mass (Li^7, Be^7, Be^8, and B^8) by the following chains:

$$\left.\begin{array}{l} He^3 + He^4 \rightarrow Be^7 + \gamma \\ Be^7 + e^- \rightarrow Li^7 + \nu \\ Li^7 + H^1 \rightarrow Be^8 + \gamma \\ Be^8 \rightarrow 2\,He^4 \end{array}\right\} \quad \text{p-p II chain} \qquad (9)$$

The latter reaction, $Be^8 \rightarrow 2\,He^4$, takes place spontaneously because Be^8 is an unstable nucleus with a half-life of only 2.6×10^{-16} sec. Instead of capturing a free electron, the Be^7 in the first step of (9) can collide with a proton to form boron:

$$\left.\begin{array}{l} Be^7 + H^1 \rightarrow B^8 + \gamma \\ B^8 \rightarrow Be^8 + e^+ + \nu \\ Be^8 \rightarrow 2\,He^4 \end{array}\right\} \quad \text{p-p III} \qquad (10)$$

The slowest step in the p-p chain is the initial one, $2\,H^1 \rightarrow D^2 + e^+ + \nu$, and this rate limiting step controls the whole chain.

It is possible to set up a group of linear differential equations for the time-variation of the abundances of each species in the p-p chains above.

These equations resemble those of radioactive decay equations, and their solutions are in the form of a series of exponential buildup and decay terms for the individual species. These solutions of the abundance equations are useful in dealing with rapid phases of stellar evolution that take place on time scales too short to allow equilibrium to be established among nuclear species. Or the equations can be used to predict equilibrium abundances among the various nuclear species during slow stages of evolution. These equilibrium abundances are inversely proportional to reaction rates involving the species. Nuclei that have slow-reaction rates will be abundant while those that react rapidly will have correspondingly low concentrations, and these relative abundances are important in establishing the buildup of heavier elements by the various modes of nucleosynthesis.

A star containing only H and He (e.g., extreme Population II stars) must operate on the p-p chains, but there is another option available for stars with an admixture of heavier elements, specifically C and N. In this case C^{12} and N^{14} serve as catalysts in a double cycle of reactions by which $4\ H^1 \rightarrow He^4$ as before. The basic CNO cycle is as follows:

$$
\begin{aligned}
C^{12} + H^1 &\rightarrow N^{13} + \gamma \\
N^{13} &\rightarrow C^{13} + e^+ + \nu \\
C^{13} + H^1 &\rightarrow N^{14} + \gamma \\
N^{14} + H^1 &\rightarrow O^{15} + \gamma \\
O^{15} &\rightarrow N^{15} + e^+ + \nu \\
N^{15} + H^1 &\rightarrow He^4 + C^{12} \\
&\quad \text{or } O^{16} + \gamma \\
O^{16} + H^1 &\rightarrow F^{17} + \gamma \\
F^{17} &\rightarrow O^{17} + e^+ + \nu \\
O^{17} + H^1 &\rightarrow He^4 + N^{14} \\
&\quad \text{or } F^{18} + \gamma
\end{aligned}
\qquad (11)
$$

The rate limiting reaction is N^{14} (p, γ) O^{15} so N^{14} eventually becomes the most abundant species in the bi-cycle. The initial C^{12} and C^{13} nuclei become converted in time mostly into N^{14}.

Besides the primary fusion reactions contained in the p-p chain and CNO bi-cycle there are other reactions that fuse other nuclei than H^1 to produce significant contributions to stellar energy generation under certain conditions. The first of these to be considered is the fusion of three alpha particles (He^4 nuclei) to form a C^{12} nucleus. The straightforward reaction

$$3He^4 \rightarrow C^{12} \qquad (12)$$

is a very unlikely reaction because it would involve a 3-body collision, moreover one in which two Coulomb barrier tunnellings must occur simultaneously. The probability for such an event even under conditions at the center of a star is so small as to rule out this reaction as a viable energy source.

There is, however, a way for the triple alpha (3α) reaction to go in two steps. The first makes a Be^8 nucleus via the inverse of the decay reaction for this nucleus; namely

$$He^4 + He^4 \rightleftarrows Be^8 \qquad (13)$$

As mentioned earlier, the Be^8 lasts only a very short time since its decay lifetime is 2.6×10^{-16} sec. In a hot, dense He^4 gas deep in a stellar interior this short lifetime is sufficient to give a reasonable chance that the Be^8 will be struck by another He^4 before it decays, yielding C^{12} by the reaction

$$Be^8 + He^4 \rightleftarrows C^{12} + 2\gamma \qquad (14)$$

Even this step would not produce appreciable carbon were it not for the fact that there is an excited state of the C^{12} whose energy level is near the combined mass-energy of Be^{8} + He^{4}. Even so, the excited C^{12} state decays back into Be^{8} + He^{4} most of the time and into a stable C^{12} state only occasionally. The large Coulomb barrier for both (α,α) and (Be^{8},α) collisions leads to a very strong temperature dependence of the overall reaction rate for the triple alpha process. The energy production rate for this process can be represented approximately by a T^{n} dependence where n varies with temperature but ranges between 30 and 40 for temperatures between 10^{8} and 10^{9}°K where this process is an important energy generator.

From the standpoint of nucleosynthesis the 3α reaction is very important for it is the bridge from very light element synthesis to the production of heavier elements. It is a hard bridge to cross, but one that can generate substantial amounts of C^{12} in late evolutionary stages of massive stars. C^{12} is the fifth most abundant nucleus in the solar system and presumeably also in the Universe, testimony to the operation of the 3α process despite the rather unlikely prospects for its success.

The third most abundant element is O^{16}, formed in part by the CNO cycle and in part by the direct capture of an alpha particle by C^{12} via the reaction

$$C^{12} + He^{4} \rightarrow O^{16} + \gamma \qquad (15)$$

In a similar way we can account for Ne^{20} synthesis,

$$O^{16} + He^{4} \rightarrow Ne^{20} + \gamma \qquad (16)$$

Under conditions favorable to this reaction, however, the Ne^{20} rapidly picks up an alpha particle to form Mg^{24} leaving a low abundance of Ne^{20}. This low abundance of Ne^{20} predicted on the basis of alpha capture does not agree with the observed natural abundance of Ne^{20} which puts it fourth in order behind O and ahead of C. In Population I stars Mg^{24} is eighth in order of abundance, well below Ne^{20} in contrast with predictions from alpha processes. The order of observed abundances in Pop. I stars is

H, He, O, Ne, C, N, Si, Mg, S, Ar, Fe, Na, Cl, Al, Ca, F, Ni (17)

Clearly the observed Ne abundance must be accounted for by processes other than (16) above.

The He burning does not involve temperatures high enough to build elements much beyond Mg. The Coulomb barrier for alpha capture becomes too great and the alpha capture rate decreases rapidly so relatively little synthesis of heavier elements can take place before all the He is exhausted in the core of a star. The star's central temperature will continue to rise as a result of gravitational contraction of its inert C^{12} core formed as He → C, and if the star is massive enough, the core temperature may reach $\approx 6 \times 10^{8\,o}K$ whereupon C^{12} begins to combine with itself in a carbon-burning phase

$$\begin{aligned} C^{12} + C^{12} &\rightarrow Mg^{24} + \gamma \\ &\rightarrow Na^{23} + H^{1} \\ &\rightarrow Ne^{24} + He^{4} \\ &\rightarrow Mg^{23} + n \\ &\rightarrow O^{16} + 2\,He^{4} \end{aligned} \qquad (18)$$

One of the reactions above yields free neutrons, and there is another important source of these particles through the reaction chain

$$\begin{aligned} &C^{12} + H^{1} \rightarrow N^{13} \rightarrow C^{13} + e^{+} + \nu \\ &C^{13} + He^{4} \rightarrow O^{16} + n \end{aligned} \qquad (19)$$

The proton that is captured by C^{12} to form N^{13} in the first step comes from the second line of equation (18).

The free neutrons liberated in these reactions are available for capture by virtually everything present in the gas, and this capture is relatively easy because there is no Coulomb barrier between a neutron which is a neutral particle and an ion. The neutrons are therefore available to synthesize a variety of heavier elements.

The alpha particles liberated in two of the C-burning steps can be captured by C^{12}, O^{16}, Ne^{20}, and Mg^{24}, and the temperature is now high enough to make these captures go at a sensible rate even with Mg^{24} to form Si^{28}.

Once all the carbon has been consumed, the central temperature of the star will rise again due to gravitational contraction, and if the star is massive enough for this temperature to reach $\approx 10^{9}$°K, oxygen burning will begin:

$$\begin{aligned} O^{16} + O^{16} &\rightarrow S^{32} + \gamma \\ &\rightarrow P^{31} + H^{1} \\ &\rightarrow S^{31} + n \\ &\rightarrow Si^{28} + He^{4} \\ &\rightarrow Mg^{24} + He^{4} \end{aligned} \qquad (20)$$

This reaction series is very much like that for C-burning, and again there is one direct neutron producing reaction.

At temperatures between 6×10^{8} and 10^{9}°K where carbon and oxygen burning take place, the thermal photons have enough energy to cause disintegration of many of the nuclei that have been built up by the reactions described above. This photodisintegration is analogous to thermal ionization of gases such as H which takes place at $\approx 10^{4}$°K. Because nuclei are bound together much more tightly (by a factor of ≈ 10^5) than electrons are bound to atoms, it follows that nuclear photodisintegration should take place at much higher temperatures, of the order of 10^9°K. For example, Ne^{20} whose abundance we have noted already poses a problem is further broken down by photodisintegration at temperatures above 10^9°K:

$$\gamma + Ne^{20} \rightarrow O^{16} + He^4 \tag{21}$$

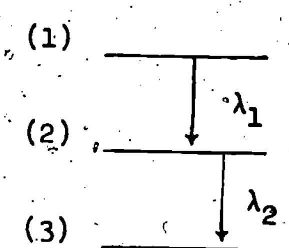
This is just the inverse of the alpha capture by O^{16}, and in general photodisintegrations are the inverse of earlier building-up reactions that yield a gamma ray, i.e., the reactions $A + \alpha \rightleftarrows (A + 4) + \gamma$ goes both ways at high enough temperatures.

At still higher temperatures, beyond about 1.3×10^9 °K where oxygen burning is complete, the nuclear reactions are mostly of a rearrangement type in which the existing nuclei are photodisintegrated, and the products of this disintegration are quickly captured by another nucleus. In such reactions there is a trend toward equilibrium in which mainly the most stable nuclei survive. There is a broad maximum in the nuclear stability curve, the curve of binding energy, in the vicinity of Fe^{56}, so equilibrium favors the formation of nuclei in the vicinity of the iron peak.

The rearrangement reactions begin to be important, i.e., to occur rapidly, at about 3×10^9 °K, and at these temperatures Si^{28} is the predominant nucleus remaining since it is more tightly bound than the other products of oxygen burning which have largely been photodisintegrated by this time. Accordingly, the rearrangement reactions leading to iron peak nuclei are often considered to begin with a predominantly Si^{28} gas. A great variety of reactions proceeds simultaneously with the result that the details become very difficult to follow. However, the overall trend toward the iron peak elements is clear. These reactions tending toward equilibrium concentrations of the elements are referred to as e- processes in the parlance of nucleosynthesis.

In order for equilibrium to be approached, it is necessary for the various rearrangement reactions to proceed faster than the elements so formed are able to undergo beta decay, and this is why the e- process requires temperatures above 3 x 10^9 degrees. At slightly lower temperatures the β-decay processes are as rapid as the photodisintegrations and lead to other than an equilibrium mixture of nuclei. This non-equilibrium mixture is very difficult to calculate in contrast with the equilibrium abundances which are easily calculated by methods of statistical mechanics from relative binding energies of the nuclei. These binding energies are accurately known from laboratory measurements.

The calculation of equilibrium abundances follows the same procedure as calculation of parent-daughter ratios in the secular equilibrium of a natural radioactive decay chain. The method can be illustrated by a simple, three element chain:



Each line in this figure represents one element designated by numbers 1, 2, 3. The differential equation for the decay of element (1) is simply

$$\frac{dN_1}{dt} = -\lambda_1 N_1$$

where λ_1 is its decay rate. For element (2) the equation is a little more complicated since $N_2(t)$ is increasing as a result of decays of N_1 which convert (1) into (2), but N_2 is also being depleted by decay into (3). Thus

$$\frac{dN_2}{dt} = \lambda_1 N_1 - \lambda_2 N_2$$

Finally, N_3 is built up entirely by decay of (2), so

$$\frac{dN_3}{dt} = \lambda_2 N_2$$

These three coupled differential equations can be solved by rather simple methods, but in a much more complex scheme such as occurs in a star in which there are very many elements and several interconnected buildup and decay chains the equations become formidably complex. In the equilibrium case in which all time variations $\frac{dN_i}{dt}$ have become zero, the equations reduce to a set of simultaneous algebraic equations rather than differential equations. In the above example, these reduce simply to $\lambda_1 N_1 - \lambda_2 N_2 = 0$ or $N_2/N_1 = \lambda_1/\lambda_2$. Thus the equilibrium ratio of N_2 to N_1 is given simply by the ratio λ_1/λ_2 of decay constants. Even with complex stellar buildup schemes the equilibrium ratios can be obtained as ratios of reaction rates without solving the differential equations.

The production of neutrons becomes copious in the temperature range above 10^9 °K by processes already mentioned, and these are available to build the neutron-rich isotopes of elements lighter than the iron peak. They can also build elements heavier than the iron group clear up to B_i^{209}. The next heavier element, P_o^{210}, is unstable and decays by α-decay as rapidly as it forms in the relatively slow neutron addition processes that go on in evolutionary stages of normal stars. These neutron capture processes that take place in late red giant stars have been termed s-processes ("slow" neutron processes). Detailed calculations based upon neutron capture cross sections measured in the laboratory (e.g. Oak Ridge) with 25 kev, neutrons have been rather successful in predicting observed relative abundances of elements heavier than Fe^{56} up to Tl^{203} and even beyond to Bi^{209} with somewhat less accuracy.

At still higher temperatures ($> 5 \times 10^9$ °K) neutron fluxes are greatly increased, and neutron capture rates greatly exceed β-decay and α-decay rates permitting the buildup of unstable elements heavier than Bi^{209} all the way to Cf^{256} and perhaps beyond, though such high atomic weight elements are very unstable and decay quickly. Not only is high temperature but also high density required to build elements above atomic weight 209. The key to production of these elements is rapidity of formation to overcome the competition of rapid decay rates, and we have seen that the reaction rates are proportional to density so high densities are required. In practice, densities above 10^8 gms/cm^3 are generally needed for the rapid neutron capture processes to be successful. These are generally called r-processes, and the temperatures and densities that make them possible as element builders are thought to occur only in the explosion of a supernova. Details of these catacysmic events will be treated in another minitext of this series entitled "Explosive Astrophysics". It will be sufficient here to say that within the order of 15 minutes, during which a typical supernova is believed to collapse and explode, the conditions are right for the r- processes to take place, and in this short period the synthesis of all elements beyond Bi^{209} occurs. In addition significant amounts of many elements between Ge and Bi are also produced by r-processes. For example, from 50 to 80% of the solar system gold abundance is believed to have been produced by r-processes. This means that most of the gold in one's dental fillings, or in the rings on one's finger was created in the interior of a supernova. Such explosions also solve the problem of how, once a heavy element is created deep within a star, it manages to escape into the interstellar regions of space. There it may later be reassembled by gravity into another star to give spectroscopic evidence of its presence, this time in the outer layers of the star, or in the case of the Sun, in its surrounding planets that were formed from the same reprocessed interstellar matter.

Bibliography

(1) G. Burbidge, E. M. Burbidge, W. A. Fowler, and F. Hoyle, Rev. Mod. Phys. 29, 547 (1957) Commonly referred to as B^2FH, this classic paper is the "bible" of modern nucleosynthesis. An excellent, more recent textbook presentation at the graduate level is

(2) D. D. Clayton, "Principles of Stellar Evolution and Nucleosynthesis", McGraw-Hill, N. Y. (1968).

A very readable and much briefer treatment can be found in

(3) M. Harwit, "Astrophysical Concepts", John Wiley & Sons, N. Y. (1973), pp. 329-355.

STELLAR EVOLUTION

I. Introduction

A star loses energy, so it must change with time. The equilibrium structure of a star is uniquely determined by its mass and composition.(1) As it radiates away energy, it must lose mass ($E = mc^2$), and this effect alone might be expected to cause a change in the star. In very luminous stars mass loss both by radiation and by stellar wind can have some noticeable effects, but in stars like the sun the mass change is very small. The sun, for example, loses less than 10^{-4} of its mass in a billion years through radiation, hardly enough to produce an appreciable change in structure.

In the energy producing core of a star a more subtle but more influential mass loss occurs. There hydrogen is being converted into heavier elements, He, C, N, O, and this causes the molecular weight of the core to increase. This increase is substantial. Suppose a star begins at "zero age" on the main sequence with a typical composition for solar type stars: $X = .72$, $Y = .26$, $Z = .02$ giving $\mu = 0.608$.(2) After sufficient time has

(1) This statement is known as the Vogt-Russell theorem and holds for nearly all types of stars.

(2) The notation here is conventional in astrophysics. X and Y are respectively the fractional compositions by weight of hydrogen and helium. Z is the weight fraction of all elements heavier than helium. Since these are fractional compositions, $X + Y + Z = 1$.

The "molecular weight" of the gas, μ, is the mean weight per particle. Since the gas in the interior of a star is fully ionized, the electrons freed by ionization must be counted among the particles in determining μ. The light weight electrons dilute the "gas" inside the star, and its effective molecular weight is considerably smaller than the usual values we are accustomed to in chemistry. Ionized hydrogen, for example, has $\mu = 0.5$ instead of its un-ionized value of 1 familiar in chemistry. Details of the derivation of the formula for μ of an ionized gas can be found in the minitext on Stellar Structure or in any standard text such as M. Schwarzschild, "Structure and Evolution of the Stars", Dover, (1966). The result is

$$\mu = \frac{1}{2X + \frac{3Y}{4} + \frac{Z}{2}}$$

passed ($\approx 10^{10}$ years for a star like the sun) most of the hydrogen in the core has been changed into helium, and the core composition might typically be $X = .02$, $Y = .95$, $Z = .03$ giving $\mu = 1.303$, more than double its initial value. Such a change in μ must cause a large change in structure, and obviously the evolved structure is quite different from the initially homogeneous one. The "old" star has an almost pure He core whereas it began as mostly H. Other consequences of the hydrogen burning follow, too, as will be seen.

We shall trace the evolution of stars of several different masses and two widely different compositions from the initial stages of formation from interstellar gas and dust, through the main stage of life spent on the Main Sequence and on to the various fates that can eventually befall stars as they grow old and exhaust their central supplies of nuclear fuel, ultimately collapsing under the grip of gravity.

A star travels through life by a succession of alternating energy production mechanisms. First, gravity causes contraction of a gassy, dusty interstellar cloud, raising its central density and temperature in the process until both are high enough to begin hydrogen fusion into helium. Then nuclear energy production takes over the job of supplying the energy the star radiates away. Since this energy source is both efficient and well supplied with fuel, it operates for a relatively long time, but not forever. The time eventually comes when hydrogen in the core is used up, and the star turns again to gravity to supply its energy requirements. The core contracts, and in so doing it heats up sufficiently to ignite fusion in a shell of H around it. This shell-burning source then takes over energy production, and the star once again relies upon nuclear energy. Gravity continues to squeeze the constantly growing helium core, and eventually, if the star is massive enough, this core begins to "burn" He to form C^{12} by the triple alpha process:

$$3\ He^4 \rightarrow C^{12}$$

In this way a new source of nuclear energy is added. Eventually the He in the core is all converted into carbon, and this nuclear energy source ceases. Gravity once more takes over, squeezing and heating the core further until, if the star is massive enough, even the carbon may begin to fuse into several products such as $Ne^{20} + He^4$, $Na^{23} + H$, and occasionally $Mg^{23} + n$, an important source of neutrons for building heavier elements. A helium burning shell will also form before the core begins burning carbon.

If the star is not heavy enough to proceed to the next possible core burning stage, it will still have one or more shell burning sources of energy, and gravity will continue to operate on the core until it becomes degenerate, whereupon contraction virtually ceases. Degeneracy is a consequence of the Pauli Exclusion Principle which governs the occupation of available energy states in a dense collection of interacting spin ½ particles such as electrons. Ordinary metals exhibit the features of a degenerate "electron gas" that are of importance in degenerate stellar interiors, namely high thermal conductivity and low compressibility. These metallic characteristics of a degenerate core will be of importance later in understanding certain phases of stellar evolution.

Our knowledge of stellar evolution rests upon two foundations. First, there are observations of stars from which such properties as mass surface temperature, surface composition, radius, and rate of radiation of energy can be determined. A great deal of effort and ingenuity by observational astronomers has gone into the gathering and interpretation of the data on which our knowledge about the mass, luminosity, temperature, radius, and composition (M, L, T, R, μ) of the stars depends. Much of this observational information is summarized graphically in the two chief diagrams of astronomy, the Hertzsprung-Russell (H-R) diagram and the Mass-Luminosity (M-L) diagram which show what values of M, L, T, μ are exhibited by real stars.

The second foundation of stellar evolution rests upon the branch of astrophysics that deals with stellar structure. It is the task of this science to reproduce by calculation from first principles stellar models that correspond in observable properties (M, L, T, R, μ) with real stars. Since real stars fall into well defined and often quite narrow regions of the H-R and M-L diagrams, it is evident that nature has placed tight restrictions upon the range of possible structures of a star. To put it another way, we might say that the laws of physics select only those stellar structures that conform to observed stars, and it is the task of the science of stellar structure to learn how properly to apply the laws of physics to yield realistic models of stars.

Among the great numbers of observable stars are to be found a wide gamut of the properties M, L, T, R, μ plus one additional variable: τ, or "age". That is, we are surely observing stars with a great range of ages from very young, newly formed stars, perhaps less than 10^5 years old, to those more than 10^{10} years old, comparable to the age of the Universe. All of these stars lie in well defined regions of the H-R diagram, and it is one of the tasks of the study of stellar evolution to discover how the observable properties of a star change with time and how a star moves in the H-R diagram as it ages.

There is clearly a very close connection between the study of stellar structure and the study of stellar evolution. The procedure to be followed in discovering how stars evolve is straightforward. One starts with a model of a real star lying somewhere on the H-R diagram, usually on the main sequence as a beginning. The internal structure of this star together with a knowledge of its luminosity or rate of energy loss permits calculation of the rate at which the composition changes inside the star where nuclear transmutations convert one kind of nucleus into another (e.g. H → He) and the rate at which gravity squeezes the interior during evolutionary phases in which gravitational energy production is significant. A new model may then be calculated based upon the altered composition, μ, of the slightly evolved star. This process is repeated many times to obtain a whole series of models evolving with time. The observable parameters, L and T, for this series will be represented by a track on the H-R diagram. These tracks are the principle means by which the results of evolutionary studies are presented, and they form the point of contact between theory and observation: the tracks must lead through regions in which real stars are found whose characteristics agree with those of the theoretical models. The time variations of internal structures of a series of evolved models give further detailed information about the changes that take place within a star as it ages, and a full appreciation of stellar evolution can only be gained by a study of such models. We will therefore supplement discussion of evolutionary tracks on the H-R diagram by description of the corresponding internal changes that take place in aging stars. The descriptions of evolution will necessarily be qualitative, but this should not obscure the fact that they are based upon carefully constructed, quantitative models in which such properties as density, temperature, etc., have been worked out as functions of radius.

II. Earliest Stages: Star Formation

The beginnings and endings of the stars remain the least understood phases of stellar evolution; middle life is the best known phase. Therefore, in discussing the initial stages in the formation of a star we are dealing with an area that remains sketchy and rather speculative. Most treatments of stellar evolution are content to begin with the obvious fact that stars <u>do</u> form and to start dealing with them in the stages just preceding arrival at the main sequence period where they remain for most of their life burning hydrogen into helium.

We shall look back further than this, however, in an attempt to outline some of the main features in the birth of a star from the gravitational contraction of a large cloud of interstellar gas and dust.

Astronomers have convincing evidence that star formation was not a one-time phase in the evolution of the Universe but is a more or less continuous process which is going on even today. They also observe that the new, young stars form only in regions of the Galaxy that have high concentrations of gas and dust. The gas, chiefly hydrogen, can be detected by radio astronomy by means of a spectral line at a frequency of 1420 MH_Z which lies in the readily accessible microwave region at a wavelength of about 21 cm. The dust consists of microscopic solid particles generally thought to be in the size range of a few thousand Ångstroms and to consist mainly of graphite and silicates, perhaps with thin coatings of ice.

Interstellar gas makes up 98-99% of the matter from which stars condense and its concentration is roughly 1 atom per cm^3. Thus the matter needed to form a star of about one solar mass ($\approx 2 \times 10^{33}$ gm) must come from an interstellar cloud whose initial radius is of order 10^{19} cm (10 light years). The temperature of this gas has been estimated to be $\approx 100^{o}K$ (from its infrared emission, among other things), and its stability may be tested by inquiring whether its gravitational attraction is strong enough to prevent its outer atoms from evaporating away. The gravitational potential energy of an atom of mass m at the edge of a cloud of mass M and radius R is $\frac{GMm}{R}$ where G is the gravitational constant (6.67×10^{-8} in cgs. units). If this is equal to the thermal energy of the peripheral atom, $\frac{3}{2}$ kT (where k is the Boltzmann constant), then the cloud will just be stable. It will collapse if

$$\frac{GM_m}{R} > \frac{3}{2} kT$$

which is the condition that the pull of gravity be stronger than the dissipating tendency of thermal motions. It is more convenient to put this inequality into a form involving the cloud density, ρ, eliminating R by the expression $\frac{4}{3} \pi R^3 \rho = M$. This gives

$$\rho \left(\frac{M}{M_\odot}\right)^2 \gtrsim \left(\frac{kT}{mG}\right)^3 \simeq 4 \times 10^{-22} T^3 \text{ gms/cm}^{-3} \qquad (1)$$

$$\text{or} \quad \rho > 4 \times 10^{-22} T^3 \left(\frac{M_\odot}{M}\right)^2 \text{ gms/cm}^3 \text{ for contraction} \qquad (1a)$$

It follows from this relation that a cloud of 1 solar mass ($M = M_\odot$) at 100° k will be stable against thermal dissipation if its initial density is greater than about 4×10^{-16} gm cm^{-3}. How does this compare with the density of the interstellar cloud containing approximately one atom per cm^3? If the atom is hydrogen, the cloud density will be $(6.03 \times 10^{23})^{-1}$ or 1.67×10^{-24} gm cm^{-3}, about a factor 10^6 too low to condense. Such a cloud will dissipate unless compressed by some external force, and the compression, if it occurs, must raise the density by a factor $\simeq 2 \times 10^6$ without increasing its temperature. If T increases during compression, Eqn (1) tells us that the compression factor must be even greater than 2×10^6, but this is likely to produce an even greater rise in temperature and is likely to be a losing proposition. It might be argued that if the compression were slow enough, the gas cloud might be able to radiate away most of its excess thermal energy and remain near 100° k. This process, however, would require a long time, of the order of 10 million years according to calculations, and it is difficult to imagine an effective compressive mechanism (remember the compression factor must be $\gtrsim 10^6$) that operates so slowly.

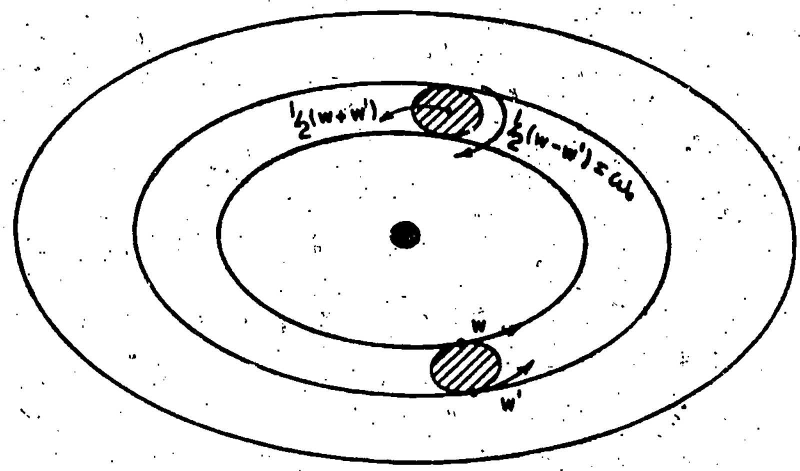
We are forced to the conclusion that a star such as the Sun probably did not condense neatly and efficiently from a cloud of about one solar mass. Instead, it most likely condensed from a much larger cloud together with a shower of other stars, probably with quite low efficiency leaving behind most of its parent gas cloud uncondensed into stars. Observations on the sizes of interstellar clouds that are fairly well defined in our part of the Galaxy indicate a range of several tens to perhaps a hundred light years in size ($\gtrsim 10^{20}$ cm) with hydrogen particle densities already up by a factor 10 over the average interstellar value of 1 per cm^3. Such clouds may contain several hundred solar masses and are often observed to be associated with groups of young stars in accordance with the suggestion that stars rarely if ever form singly.

A 100 $M_\odot$ cloud at 100°K needs a density greater than 4×10^{-20} gm cm^{-3} to be stable against thermal dissipation, according to Eqn. (1), and a cloud of 10^3 $M_\odot$ would need $\rho > 4 \times 10^{-22}$. Compared with $\rho = 1.6 \times 10^{-23}$ gm cm^{-3} for a particle density of 10 atoms/cm^3, even clouds this large are unstable, but not by so large a factor as before. More modest compression by external forces (radiation pressure from nearby stars, collisions with other gas clouds, magnetic fields) now becomes conceivable as a means of raising the density to the critical point that will initiate gravitational contraction and subsequent star formation.

Once gravitational contraction begins, all is not simple because the temperature tends to rise, and this rapidly increases the tendency of the cloud to dissipate again (note the T^3 factor in Eqn. (1)). As long as the cloud remains transparent to its own thermal radiation, there is a chance for it to radiate away energy fast enough to keep itself cool enough to prevent the contraction from stopping short of star formation. The presence of dust grains is especially valuable in converting gas atom thermal energy into thermal radiation: the gas atoms collide with the grains, heating them, and the grains radiate at long wavelengths into space like microscopic blackbody radiators. The gas is quite transparent to the far infrared radiation of the grains as long as temperatures remain low, and the whole mixture of gas and dust appears able to radiate well enough to let contraction continue once it begins. Many details are still not clear and cannot be pursued here.

As the cloud contracts it heats up, and it also spins faster since its angular momentum is conserved. Let us first consider the consequences of angular momentum conservation because it poses a problem in star formation every bit as formidable as the contraction vs. thermal dissipation problem outlined above. Consider a gas cloud with a radius 10^{20} cm and a mass of $100\ M_{\odot} \simeq 2 \times 10^{35}$ grams. If it is a uniform spherical cloud, its initial angular momentum can be calculated by elementary principles to be $\frac{2}{5} MR^2$ (its moment of inertia) times its initial angular velocity ω_o.

The cloud will have, on the average, an initial angular velocity imparted by the shear forces that arise from differential rotation about the center of the Galaxy. This rotation of the cloud arises from the fact that parts of the cloud closer to the center of the Galaxy have higher orbital speeds than those farther away (Kepler's 3rd Law says that the orbital period varies as the 3/2 power of the mean radius of the orbit). The figure below shows how differential rotation arises from the fact that $\omega > \omega'$.



Matter close to the Galactic center has higher orbital speed than at the periphery. $w > w'$. This phenomenon is called *differential rotation.*

Figure 1

In our region of the Galaxy the differential angular velocity, $\omega_0 \approx 3 \times 10^{-16}$ sec^{-1}; therefore a cloud of the size under consideration would have angular momentum $\frac{2}{5} MR^2 \omega_0 \approx 2 \times 10^{59}$ gm cm^2 sec^{-1}. Compare this with the angular momentum of the solar system, approximately 2×10^{50} gm cm^2 sec^{-1}. The two values differ by a factor 10^7 when allowance is made for the fact that the cloud contains 100 solar masses while the Sun contains only one. In short the 100 $M_\odot$ cloud must lose all but 10^{-7} of its initial angular momentum if it is to give birth to 100 well behaved stars. Put another way, conservation of angular momentum requires that $R^2\omega$ remain constant as the cloud contracts. A factor $R\omega$ of this quantity is the equatoral surface velocity of the rotating cloud, so Rv_s must be constant. Initially $v_s \approx 3 \times 10^4$ cm/sec when $R \approx 10^{20}$ cm By the time R has decreased by a factor of 10^6, to 10^{14} cm, v_s has arisen to the speed of light! And 10^{14} cm is still quite large, about 1000 times the radius of a star like the Sun.

Clearly the contracting cloud must reduce its angular momentum somehow if it is to succeed in forming stars. At present we don't know how it accomplishes this. Although there have been many speculative suggestions over the years, none has been supported by fully convincing arguments. One of the more plausible suggestions is that the cloud forms binary stars preferentially, and these are able to contain much more angular momentum than single stars. Observations confirm that a majority of stars (≈ 60%) do indeed belong to binary or multiple star systems, so formation of binaries does seem to be the norm.

In order to discuss the relative angular-momenta of different configurations, it is convenient to deal with angular momentum per unit mass which will be designated, $\widetilde{L}$. The primordial cloud we have been discussing has $\widetilde{L} \approx 10^{24}$ cm² sec⁻¹ while the solar system has $\widetilde{L} \approx 10^{17}$ cm² sec⁻¹. A binary system consisting to two solar mass stars in circular orbits at a distance r from each other has $\widetilde{L} = \frac{r^2\omega}{4} \approx \pi r^2/2P$ where P is the orbital period of the binary pair. Kepler's 3rd Law provides another relation, $MP^2 = Kr^3$, between P and r, and this leads to $\widetilde{L} \propto P^{1/3}$ or $\widetilde{L} \propto r^{1/2}$ for a binary. Putting in the constant K and the other numerical factors and inserting a statistically weighted average period for observed binaries, we find that $\widetilde{L} \approx 2 \times 10^{19}$ cm² sec for typical binary stars, a value about 200 times greater than that for an average single star. From this kind of reasoning we see that binary formation is a much more efficient mechanism for storing angular momentum than single star formation. Even so, the primordial cloud under consideration has enough $\widetilde{L}$ to supply ≈ 10^5 binaries, far more than we believe actually form from such a cloud, on the basis of observational evidence. If only a small fraction of the cloud mass actually condenses into stars we can account for observed rates of new star formation, and the cloud retains most of its initial angular momentum. Details of the fragmentation process and what happens to halt contraction of the whole cloud once star formation begins remain obscure.

There is a third ingredient to be taken into account in the early stages of star formation, the interstellar or galactic magnetic field. This field is very small (of the order of a few times 10^{-6} gauss) in the spiral arms of the Galaxy where stars are formed. Compare this galactic field with the Earth's magnetic field of ≈ 0.6 gauss to appreciate its small size. However, even so small a field, acting as it does over vast distances, can have a very strong influence upon certain stages of star condensation. As long as the interstellar gas is cool enough to remain neutral, the field does not interact directly with it. Once the gas becomes ionized, however, the fields and the charged particles interact strongly, and the motions of the ionized gas are greatly influenced by the presence of the field. Likewise the field lines are influenced by charged particle motions and the whole complex of field plus ionized gas becomes an interdependent entity, a magnetohydrodynamic plasma. The properties of such plasmas cannot be discussed in any detail here, but we can illustrate in a simple way one serious problem that comes from the presence of the magnetic field.

Consider a protostar of about 10^{17} cm radius with sufficient density to be able to contract to form a real star of radius about 10^{11} cm, a realistic value (the Sun's radius is $\approx 7 \times 10^{10}$ cm). If the protostar of 10^{17} cm radius has an appreciable amount of ionized hydrogen, as calculations indicate it will, these ions are bound to the magnetic field lines. Sufficient ions will be formed in the cloud by cosmic rays and by ultraviolet radiation from nearby hot, blue stars. Through the mechanism of collisions between ions and neutral atoms, the latter are indirectly connected with the field. Unless the fraction of gas that is ionized is extremely small, this indirect connection between gravity and the magnetic field through the intermediary of neutral particles colliding with ions is strong enough to effectively tie gravitational contraction to the magnetic field. As the gas contracts it therefore compresses the field in the plane perpendicular to the field, as illustrated in the sketch below:

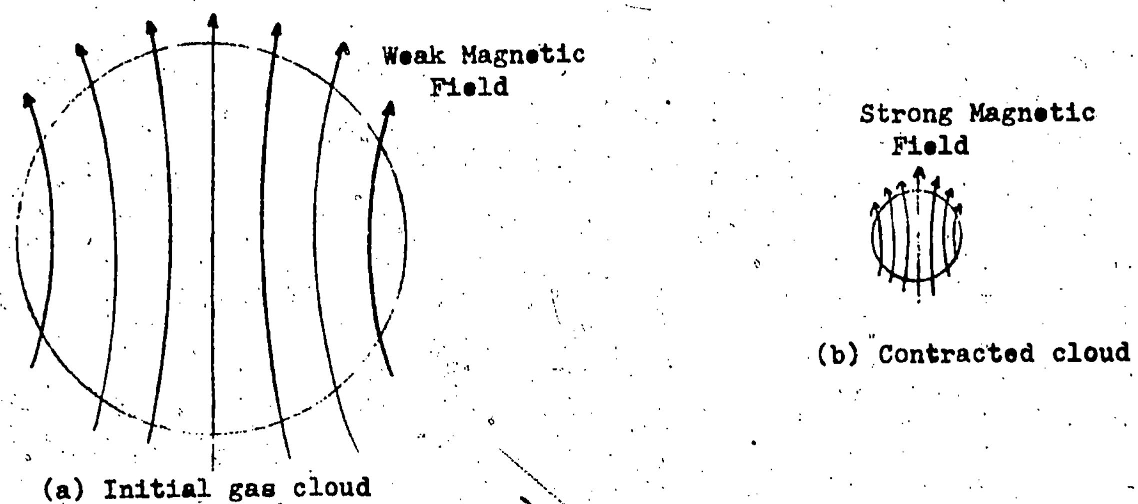


Figure 2. Compression of magnetic field lines in a contracting partially ionized gas cloud.

Conservation of magnetic flux in such a contraction leads to the conclusion that the product BR^2 is conserved where R is the radius of the cloud. Thus $B \propto R^{-2}$, and in going from $R \approx 10^{17}$ cm to $R \approx 10^{11}$ cm, B increases from initial values of a few times 10^{-6} gauss to well over 10^6 gauss. Such high fields have never been found in normal stars. The strongest fields observed have been found in certain peculiar A-type stars with values of B up to $\approx$30,000 gauss. Most stars like the Sun have fields ranging from a few gauss to a few hundred gauss. Clearly the star finds some way to rid itself of most of the flux lines that pass through its parent gas cloud, but the means by which it accomplishes this are still not understood.

We have gone to some lengths to illustrate serious difficulties that hamper our present knowledge of the earliest stages of star formation. The fact that these difficulties can be presented in terms of elementary physical principles makes it appropriate to discuss them in an introduction to stellar evolution such as this. But whatever means stars employ to overcome the barriers to their formation, they somehow manage to do so, and once they reach a certain stage of condensation, (radius $\approx 10^{13}$ cm, density $\rho \approx 10^{-6}$gm cm^{-3}, temperature $\approx 10^{3}$°K) we can follow the evolutionary history of the star with much greater confidence except during a few short lived stages late in life.

Let us first consider paths on the H-R diagram of stars from the state described above until they reach the Main Sequence. As long as the protostar obtains its radiant energy from gravitational contraction (before nuclear energy sources begin to supply its energy when it reaches the Main Sequence), there is a very useful general physical theorem, the Virial Theorem, that is very helpful in understanding the star's development. This theorem comes from considerations of the overall effect of converting gravitational potential energy, -V, into kinetic or thermal energy, K.E. The theorem says that in such conversions

$$\Delta \text{ K.E.} = \frac{1}{2}|\Delta V|$$

That is, half the potential energy released by contraction goes into kinetic energy of thermal motion (i.e. into raising the temperature of the gas); the other half leaves the star as radiation. This simple relation permits us to follow the average temperature and luminosity of the contracting star as long as it obeys the conditions for validity of the Virial Theorem. These conditions apply as long as the star is contracting slowly enough to maintain approximate hydrostatic equilibrium in which gas pressure balances the force of gravity.

There are phases of contraction during which the Virial Theorem does not hold. The first of these is encountered when the mean temperature of the gas reaches approximately 2000°K and the molecular hydrogen present dissociates into neutral atomic hydrogen. This is an endothermic reaction that absorbs thermal energy, converting it into potential energy of dissociation. That is, $H_2 \rightarrow 2H$ - energy. When dissociation occurs at a rapid rate, the star undergoes rapid contraction to supply the required dissociation energy. Its temperature remains nearly constant, but its radius decreases and its luminosity must drop, too, since $L = 4\pi R^2 \sigma T^4$. This phase, therefore, appears on the H-R diagram as a nearly vertical drop in L at nearly constant T, the segment A to C in Figure 3.

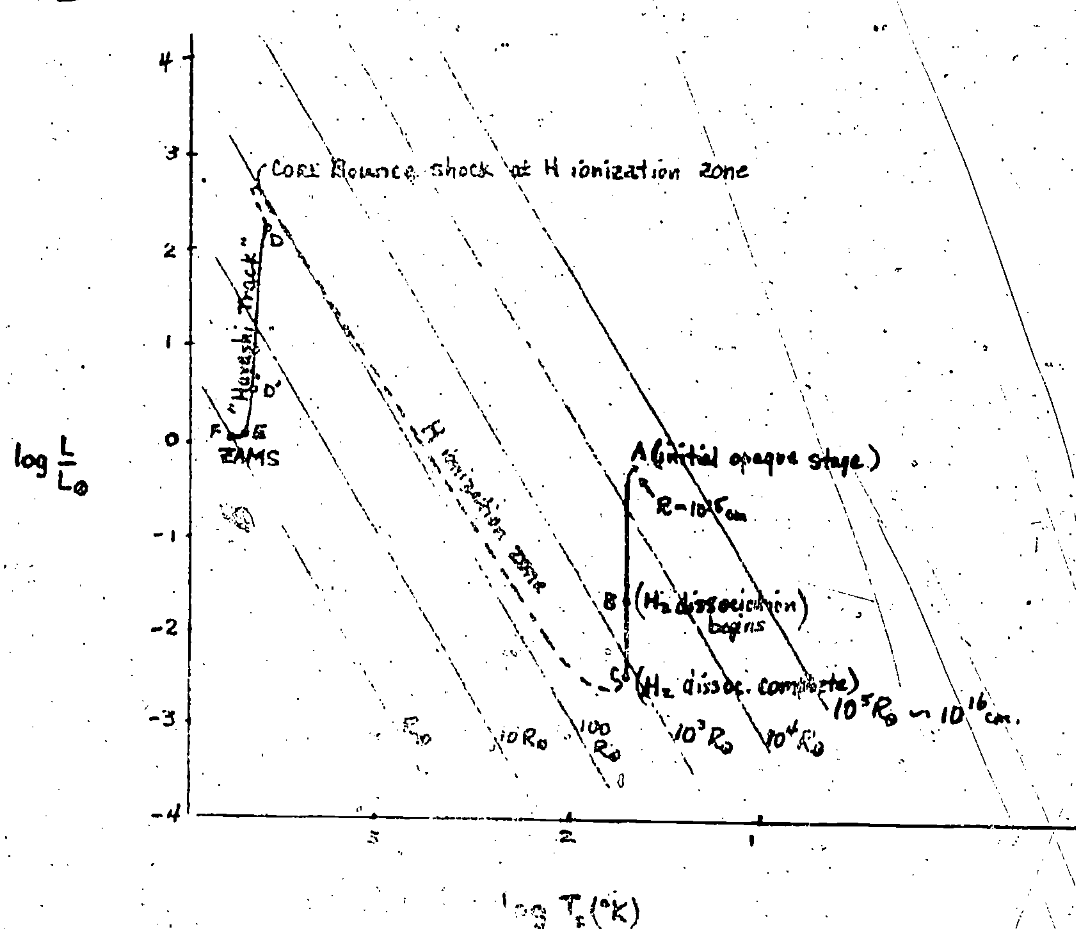


Figure 3. H-R Diagram showing early gravitational contraction stages of star formation.

This dissociation phase of the collapse begins at $R \approx 10^{15}$ cm and is complete when the radius has decreased to about 1000 solar radii ($\approx 7 \times 10^{13}$ cm). Figure 3 does not attempt to trace the collapse between the initial cloud ($R \approx 10^{19}$ cm, $T \approx 100^{\circ}K$) down to Point A. This is a dynamic free-fall process, and the cloud is transparent to the long wavelength infrared radiation (mainly at 28 microns) by which it radiates its excess energy. This phase is characterized as free-fall because the cloud is so tenuous that there is insufficient pressure to halt or even slow appreciably the gravitational pull on individual gas molecules and atoms. These fall initially inward toward the center of the cloud almost as if there were no other particles present to impede their fall. The time required for a particle to fall to the center under such idealized conditions is calculable by elementary methods. It is

$$\tau_{ff} \approx \sqrt{\frac{R^3}{GM}} \text{ sec}$$

The free-fall time for the initial gas cloud turns out to be $\approx 7 \times 10^6$ years, assuming its mass is initially $\approx 100\ M_{\odot}$ before it fragments into solar mass size stars sometime during this phase.

After dissociation is complete at C, the gas temperature can rise again as gravity contracts the star farther. The rise in temperature is aided considerably by the fact that it is now opaque; in fact, calculations show that it just becomes opaque for radiation originally at the center at A just prior to dissociation. This helps the star trap its radiation inside temporarily and speeds up the dissociation process. The interior temperature continues to rise rapidly to $\approx 9000^{\circ}K$ where ionization of atomic hydrogen takes place. This too is an endothermic reaction, and while it operates, the star is far from being in equilibrium. It continues to shrink rapidly along the dashed line (dashed because of uncertainty about details in this phase) from C to D.

The radius decreases again by about a factor 20, but this time the surface temperature of the star (the abscissa of the Figure) does not remain constant as it did before. In going from A to B the star was not yet fully opaque. At first all radiation except that originating at the center could escape without being absorbed at least once (Point A); later between A and B, as the density increased, radiation originating farther and farther from the center was just able to escape without absorption. Only at the end, at C, had the opacity risen to the point that all radiation except that from a thin layer at the surface was absorbed before escaping. Actually, at C and beyond most radiation in the interior is absorbed and reemitted many times by atoms in the stellar interior before it finally reaches the surface. This takes some time, and the surface radiation no longer reflects the instantaneous state of the interior as it did earlier when the star was still somewhat transparent. The net result is that the interior heats up very rapidly at B to the 9000°K or so required to ionize hydrogen, but the surface temperature plotted on the H-R diagram lags in time, rising until it finally reaches ≈ 9000°K at Point D. The whole process of dissociation, ionization, and final buildup of surface temperature, the trip from C to D, only takes about 20 years for a solar mass star, and this must be reckoned as virtually instantaneous on the usual time scale for stellar processes.

The little U-shaped excursion at the end of the dashed path just before Point D is the result of a minor implosion of sorts. The star collapses so rapidly that an inward radial shock wave forms, and when it reaches the now rather dense core near the center of the star, it is reflected back out, a process vividly described as a "core-bounce". The shock wave travelling back out reaches the surface where its higher temperature becomes evident for a short time as a transient temperature and luminosity spike as shown adjacent to D. This transient flare up lasts about three months in a solar mass star.

At Point D the interior of the star is turbulent as a result of the rapid heating, the shock waves, and the generally non-equilibrium nature of its preceding collapse stages. Under these conditions it transports energy from the interior to the surface mainly by convection. This is a relatively efficient form of heat transport, and it serves to stabilize the surface temperature of the star as it continues to contract toward its eventual size that it will reach on the Main Sequence. During its contraction along D to E and F at nearly constant T_s the interior temperature continues to rise rapidly, especially in the central regions. The convective, turbulent conditions in the interior insure effective mixing of the material of the star so when the central temperature eventually reaches about 2×10^6 °K at E, any deuterium, lithium, beryllium, boron, C^{13}, N^{13}, etc., are capable of undergoing nuclear reactions that destroy them. Because of the convective mixing of gas into the hot core in this phase, virtually all of these esoteric light nuclei in the star are "burned" before the star reaches the Main Sequence, and observations confirm abnormally low concentrations of these nuclei in normal stars.

As the star contracts from D to E along the nearly vertical track (the "Hayashi track" as it is called after its discoverer) it remains in quasi-hydrostatic equilibrium, a condition amenable to study by standard methods of calculation of stellar interiors. The star has finally reached a well-behaved condition that permits some insight into its internal structure, and a number of researchers have calculated detailed series of evolutionary models of stars in this pre-Main Sequence phase. These calculations show, for example, that at Point D the star begins to depart from a condition of complete turbulance and a "radiative core" develops, so-called because energy transport within it is by means of radiative transfer. That is, energy moves outward in the form of radiation that is successively absorbed and reemitted by the gas particles, now almost entirely ions and electrons at the temperatures of the core. As time goes on the radiative core grows at the expense of the convective region, and the boundary between the two gradually receeds toward the surface. Radiative transfer is a quieter but less efficient means of energy transport than convection, and it drastically changes the internal character of the star.

Figure 4 gives the results of a number of pre-MS evolutionary tracks calculated by Iben, and Table I following it lists the times in years required to reach the corresponding points on Figure 4 beginning at Point D in Figure 3, the start of the Hayashi track.

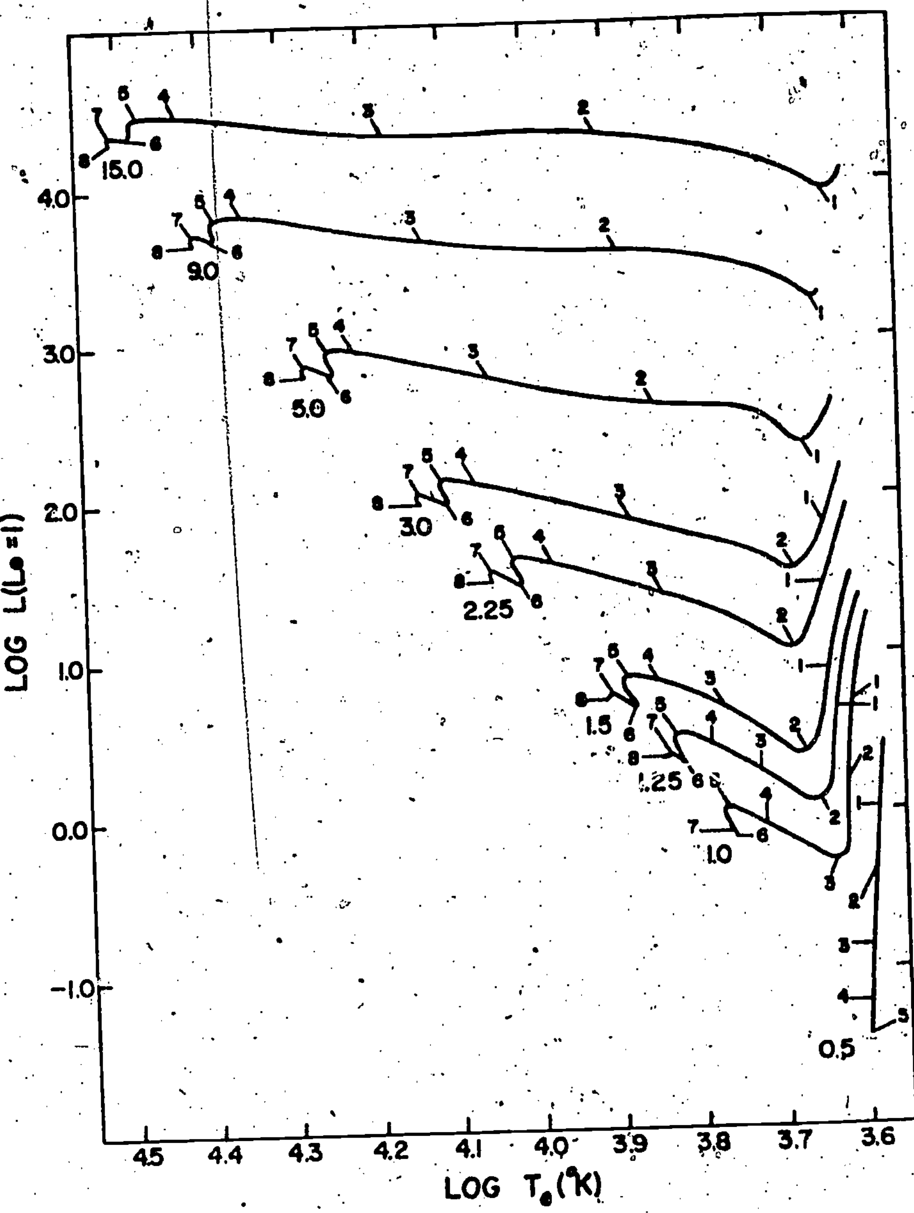


Fig. 4 Pre-MS evolutionary tracks on the H-R diagram for models of mass M (solar units) = 0.5, 1.0, 1.25, 1.5, 2.25, 3.0, 5.0, 9.0, and 15.0 (indicated alongside the appropriate tracks). Numbered points along the tracks are the points referred to in Table I

EVOLUTIONARY LIFETIMES (YEARS)†

Point	M (solar units) 15.0	9.0	5.0	3.0	2.25
1	6.740(2)	1.443(3)	2.936(4)	3.420(4)	7.862(4)
2	3.766(3)	1.473(4)	1.069(5)	2.078(5)	5.940(5)
3	9.350(3)	3.645(4)	2.001(5)	7.633(5)	1.883(6)
4	2.203(4)	6.987(4)	2.860(5)	1.135(6)	2.505(6)
5	2.657(4)	7.922(4)	3.137(5)	1.250(6)	2.818(6)
6	3.984(4)	1.019(5)	3.880(5)	1.465(6)	3.319(6)
7	4.585(4)	1.195(5)	4.559(5)	1.741(6)	3.993(6)
8	6.170(4)	1.505(5)	5.759(5)	2.514(6)	5.855(6)

Point	M 1.5	1.25	1.0	0.5
1	2.347(5)	4.508(5)	1.189(5)	3.195(5)
2	2.363(6)	3.957(6)	1.058(6)	1.786(6)
3	5.801(6)	8.800(6)	8.910(6)	8.711(6)
4	7.584(6)	1.155(7)	1.821(7)	3.092(7)
5	8.620(6)	1.404(7)	2.529(7)	1.550(8)
6	1.043(7)	1.755(7)	3.418(7)	–
7	1.339(7)	2.796(7)	5.016(7)	–
8	1.821(7)	2.954(7)	–	–

† Numbers in parentheses are the powers of 10 by which the corresponding entries are to be multiplied.

Table I

The general time scales for Hayashi tracks ($D \rightarrow E$) of stars of various masses should be noted. The Sun, for example, requires about 10^6 years to reach Point D' in Figure 3 (Point 2 in Figure 4) where development of a radiative core begins and another 8×10^6 years to reach E where this development is complete and the star turns to the left (points 3 to 5 in Figure 4) almost parallel to the M.S. The same sequence in a 3 $M_\odot$ star, on the other hand, takes only $\approx 8 \times 10^5$ years, a factor 10 less. This illustrates clearly how rapidly the evolutionary time scale speeds up as one considers more massive stars. By contrast a 0.5 $M_\odot$ star continues all the way to the M.S. on its Hayashi track taking $\approx 10^8$ years to do so. A star more massive than $\approx 3\ M_\odot$ hardly has any Hayashi track at all, and massive stars scarcely appear to experience any break in the leftward trend of their tracks beyond the onset of opacity. The rapid pre-MS evolution of heavy stars is noteworthy. For example a 15 $M_\odot$ runs through the entire gamut, 1 through 8, of Figure 4 in only 60,000 years, a factor 10^3 less than a solar mass star.

The mass-luminosity relation gives a good approximate idea about relative time scales in stellar evolution. Both theory and observation reveal that luminosity L is proportional to some power of the mass of a star. This can be written

$$L \propto M^n$$

where n varies somewhat with a star's composition and position on the H-R diagram, ranging from ≈ 2.5 for very low mass stars to over 5 for massive stars. The L-M relation can then be written in terms of solar values as

$$\frac{L}{L_\odot} = \left(\frac{M}{M_\odot}\right)^n$$

The luminosity of a star is its rate of energy loss and the time spent by a star during a certain phase of its evolution should be inversely proportional to the rate at which it loses energy (i.e., $\tau \propto L^{-1}$). This, taken together with the mass-luminosity relation leads to the expectation that the time scales for stars to run through corresponding stages of evolution should follow the relation

$$\frac{\tau}{\tau_{\odot}} \propto \left(\frac{M_{\odot}}{M}\right)^{n}$$

Examination of Table I reveals that n ranges from ≈ 1.5 to ≈ 2.5 depending somewhat upon mass but more upon the stage of evolution considered. $n \approx 2.5$ is a fair average for the trip from 1 to the Main Sequence in Figure 4.

The slightly rising, leftward movement of a star in Figure 4 represents a stage before the onset of nuclear energy production in which the star still obtains its energy by gravitational contraction with a hot, compact, radiative core. When this core becomes hot enough, at point 5, nuclear reactions are ignited and begin converting hydrogen into helium plus considerable energy. The detailed reactions are fairly complex and are discussed elsewhere in the mini-texts on Stellar Structure and Nucleosynthesis. The overall result may be summarized

$$4\,H^{1} \rightarrow He^{4} + \text{energy}.$$

The energy comes from the fact that a He^{4} nucleus is slightly less massive than 4 H^{1} nuclei. The mass difference appears as energy according to the Einstein relation $\varepsilon = \Delta mc^{2}$ where Δm is the mass difference. This energy appears initially mainly in the form of gamma rays and constitutes about 0.7% of the total mass-energy of the reacting nuclei. Thus the process is less than 1% efficient in converting mass into energy, but that low efficiency is adequate to keep the star going for a very long time, over 10 billion years for a solar type star.

As soon as nuclear fusion taken over the job of furnishing energy, the star's interior readjusts itself, contracts slightly, and its luminosity decreases a bit as it settles down onto the Main Sequence to spend the major phase of its life quietly converting hydrogen into helium in its central regions. The last points shown for each model star in Figure 4 (point 8 in most cases) represent arrival on the MS. By custom, arrival on the MS is reckoned as the true birth of the star and at this point it is considered a Zero Age Main Sequence star. The star is of course observable throughout much of its pre-MS gestation period where its luminosity is actually higher than the MS star it will become. Observationally the T Tauri and Herbig emission classes of stars are believed to consist of stars in the pre-MS region, and the Herbig-Haro objects may well be still earlier pre-MS stars in the pre-dissociation stage of Figure 3. Thus stars become visible well before their conventional birth, but at ZAMS they become much easier to observe because they spend so much time on the MS that there are a great many there at any given time for us to see. By contrast, stars move rapidly through many phases of their pre-MS period and consequently at any given time there are few present in these phases to be observed.

III. Middle Stages: Main Sequence

Stars on the MS obtain their energy by two principle series of fusion reactions. Both have the same net result, namely fusion of 4 protons into one helium nucleus. But one series of reactions, the proton-proton chain involves only light elements: He^4, He^3, D^2, H^1, and it operates effectively at relatively "low" temperatures (low for stellar interiors) between ≃5 to 25 million degrees. The energy generation rate for this process is roughly proportional to the fourth power of the temperature near the center of the star. The other process involves the normal isotope of carbon, C^{12}, in a complex series of reactions that eventually regenerates the C^{12} again, so carbon acts as a catalyst in the process. Various isotopes of nitrogen and oxygen are also involved and the process is known variously as the carbon chain, the CN chain, or the CNO bi-cycle. It begins operating at about 12 million degrees and continues up to $\approx 50 \times 10^{6}$ °K. This cycle is therefore effective over a somewhat higher temperature range than the p-p chain, and it is characterized by a much stronger temperature dependence, $\approx T^{20}$ at the low end of the range dropping to $\approx T^{14}$ near the upper end.

The large temperature dependence of the carbon cycle has an important consequence. It leads to a turbulent, convective core in stars in which this mode of energy production dominates, i.e. in stars a little more massive than the Sun with central temperatures above $\approx$ 15 million degrees. These more massive stars that operate mainly on the carbon cycle with convective cores and radiative envelopes are designated Upper Main Sequence (UMS) stars. The less massive stars which operate on the proton-proton chain with radiative cores and convective outer layers are designated Lower Main Sequence (LMS) stars. The difference in structure between these two broad types of MS stars has important consequences in the manner in which they evolve, as we shall see.

Once a star reaches the Zero Age Main Sequence (ZAMS), it begins in ernest to convert its hydrogen core into helium. Only the central 15-20% of the star is hot enough and dense enough to sustain hydrogen fusion, and it is only in this core that the chemical composition changes while the star is on the M.S. If the star is less massive than $\approx 1.1\ M_{\odot}$, the core is quiet, non-convective, and non-turbulent because of the relatively low ($\approx T^{4}$) dependence of the energy generation rate. The temperature rises to a maximum at the center of such a core, and so does the rate of energy generation. Since there is no convective turbulence, there is no mixing going on in the core and the fusion product, helium, stays where it is formed. It forms most rapidly at the center where the rate of fusion is highest, and in time there develops a small but growing subcore of pure He. Nuclear energy production ceases in this subcore since it is not hot enough to fuse He into C. It remains inert, the "ashes" of hydrogen burning, while hydrogen fusion continues in a thick shell outside the He subcore. This He core continues to grow, of course, as the hydrogen burning shell makes more helium which is added to the ever-growing He subcore. This pure He core remains at constant temperature rather like a loaf of bread in an "oven" of hydrogen burning shell. The shell continues to burn outward as a grass fire burns in a thin, ever widening circle leaving a growing, charred central area. In this fasion the He core grows until its mass becomes about 10% of the total mass of the star. Gravity begins to squeeze the core, raising its temperature significantly so it is no longer isothermal. Eventually its central temperature will reach ≈ 100 million degrees which is high enough to initiate helium fusion by the "triple-alpha" process, $3\,He^4 \rightarrow C^{12}$. This process generates nuclear energy, and the star now has two nuclear energy sources, a He burning core and a hydrogen burning shell. All of this is still deep within the star, but the effect is felt at the surface of the star where the additonal energy generated by He fusion increases both the luminosity and the size of the star while at the same time lowering its surface temperature. The star at this stage in its evolution has therefore left the main sequence and become a red giant.

Figure 5 shows post-MS evolutionary tracks of a series of stars ranging from 1 to 15 solar masses and Table II shows the time intervals between numbered points on the tracks. Point (1) is the ZAMS point for each star. The stage (1) → (2) corresponds with exhaustion of H in the core. The stage

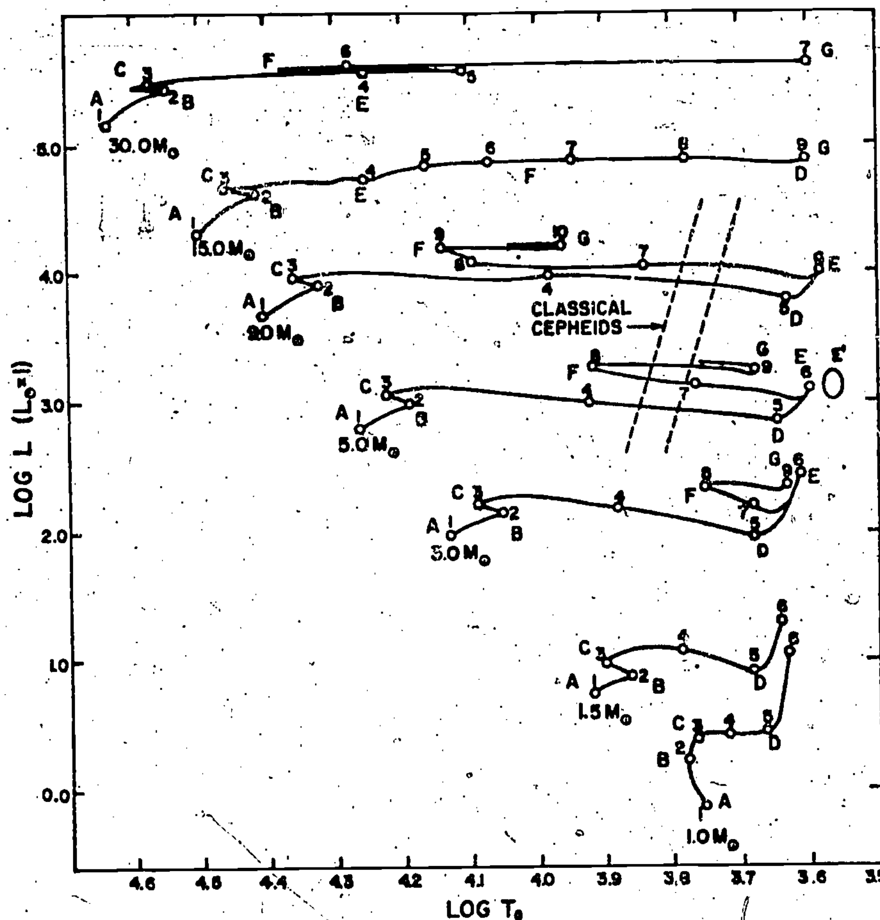


Fig. 5. Post-MS evolutionary tracks on the theoretical H-R diagram for stars having masses in the range $1.0 \leq M \leq 30.0$ and, initially, a "Population I" composition ($X = 0.708$, $Z = 0.02$ for all tracks except the $M = 30$ track; for this track, $X = 0.70$, $Z = 0.03$).

Table II. EVOLUTIONARY LIFETIMES (YEARS)†

(*Initial composition:* $X = 0.708$, $Z = 0.02$ *for* $1.0 \leq M \leq 15.0$; $X = 0.70$, $Z = 0.03$ *for* $M = 30.0$)

	M (solar units)						
Point	1.0	1.5	3.0	5.0	9.0	15.0	30.0
1	5.016(7)	1.821(7)	2.510(6)	5.760(5)	1.511(5)	6.160(4)	2 (4)
2	8.060(9)	1.567(9)	2.273(8)	6.549(7)	2.129(7)	1.023(7)	4.82(6)
3	9.705(9)	1.652(9)	2.394(8)	6.823(7)	2.190(7)	1.048(7)	4.91(6)
4	1.0236(10)	2.036(9)	2.478(8)	7.019(7)	2.208(7)	1.050(7)	4.92(6)
5	1.0446(10)	2.105(9)	2.488(8)	7.035(7)	2.213(7)	1.149(7)	4.93(6)
6	1.0875(10)	2.263(9)	2.531(8)	7.084(7)	2.214(7)	1.196(7)	5.45(6)
7	–	–	2.887(8)	7.844(7)	2.273(7)	1.210(7)	5.46(6)
8	–	–	3.095(8)	8.524(7)	2.315(7)	1.213(7)	–
9	–	–	3.262(8)	8.782(7)	2.574(7)	1.214(7)	–
10	–	–	–	–	2.623(7)	–	–

* Numbers in parentheses are the powers of ten by which the corresponding entries are to be multiplied.

(2) → (3) corresponds with development of a hydrogen burning shell source outside of the isothermal He core, and at Point (3) the core has reached the limiting mass (≃ 10% of the mass of the star) beyond which the core begins to heat rapidly under gravitational contraction, developing a hot center that ultimately ignites He fusion.

It is evident from Figure 5 that there is a qualitative difference between the shapes of the tracks (1) → (2) → (3) in a low mass star ($\lesssim 1.1\ M_\odot$) and tracks of more massive stars. The low mass stars have non-convective cores and develop H-burning shells gradually and without fuss. This produces a smooth track without breaks or kinks as the limiting isothermal core mass is reached at (3). It takes a 1 $M_\odot$ star about 9 billion years to traverse the track from (1) to (3), as Table II shows.

It is easy to estimate the time a normal star spends on the Main Sequence. Let its total mass be M gm and its luminosity be L ergs/sec. The efficiency of hydrogen burning is .007 meaning that 1 gram of hydrogen converted into He gives .007 c^2 ergs of energy. If the star stays on the Main Sequence until 10% of its mass has been converted into He, the total energy released is $7 \times 10^{-4}\ Mc^2 = 6.3 \times 10^{17}$ M ergs. If the star radiates energy at the rate L ergs/sec for a time τ on the M.S., then

$$\tau_{ms} \simeq 6.3 \times 10^{17}\left(\frac{M}{L}\right)\text{seconds}$$

Applied to the Sun ($L = 4 \times 10^{33}$ ergs/sec, $M = 2 \times 10^{33}$ gm) one gets τ ≃ 3×10^{17} sec = 10^{10} years which is just slightly more than the 9 billion years calculated for the Sun to evolve from points (1) to (3) in Figure 5. For other stars the equation above can be put into the form of a ratio to solar values:

$$\tau_{ms} \simeq 10^{10}\left(\frac{M}{M_\odot}\right)\left(\frac{L_\odot}{L}\right)\text{ years.}$$

It is possible to go one step further by using an approximate mass-luminosity relation, $L \propto M^4$ which is a reasonably good approximation over a large part of the Main Sequence. With this relation, the lifetime of a star on the M.S. becomes

$$\tau_{ms} \simeq 10^{10} \left(\frac{M_{\odot}}{M}\right)^3 \text{ years.}$$

We should expect a 9 $M_{\odot}$ star to have $\tau_{ms} \simeq 1.4 \times 10^7$ years which is not far from value of 2.1 x 10^7 years for a star of this mass from Table II. It appears, therefore, that this very simple theory gives a good account of the time spent by a star on the Main Sequence.

Heavier stars than about 1.5 $M_{\odot}$ have turbulent, convective cores initially since they operate more or less on the CNO cycle with a very sensitive temperature dependence $\simeq T^{20}$. The core turbulence keeps the core well mixed and prevents the smooth growth of an inert He subcore beginning at the center. Instead, as H-burning goes on, the entire core (15-20% of the stellar radius) uniformly changes composition, becoming more and more rich in He as the H content declines. This process must eventually result in a core too poor in H to burn any more; the ashes have diluted the fuel till the fire goes out, or nearly so. At Point (2) in Figure 5 such a situation has occurred. As the nuclear energy source begins to fail, gravity steps in to take over energy production by contracting the core. This raises the temperature of the core and manifests itself by an increase in both luminosity and surface temperature of the star, causing it to move from (2) to (3) as shown in Figure 5. During this stage of evolution a massive star contracts slightly causing the track (2) → (3) to lie in the direction shown (toward higher T, L). The sharp kink at (2) reflects the relative suddenness with which the entire core runs out of hydrogen. The low mass stars have no such sudden fuel exhaustion and their cores remain isothermal between (2) and (3) while the H-burning shells continue to supply nuclear energy. As a consequence, the radius grows smoothly in a less massive star, its luminosity increases, and its surface temperature stays nearly constant, as the Figure shows for the case of a star like the Sun.

In the UMS stars the core contraction phase (2) → (3) is relatively rapid, requiring on the average only 2-5% of the time from (1) → (2). In these heavy stars, the core becomes hot enough at about (3) to ignite H-burning in the surrounding shell. The exhausted core may or may not become isothermal for a time after the H-burning shell has been ignited. It all depends upon whether or not the He core is more massive than 10% of the star's mass. For stars more massive that ≈ 6 $M_{\odot}$ an isothermal core never develops, and the gravitational contribution to energy production continues to be effective, raising the core central temperature to the triple alpha ignition point straightaway without a pause for the core to gain more He mass from the H-burning shell.

There is one more feature of LMS stars to be mentioned. For stars of < 1.3 $M_{\odot}$ the isothermal core is dense enough and cool enough to become degenerate even though its temperature may be 15 million degrees! As described earlier, degeneracy is a dense, almost incompressible metal-like state of high thermal conductivity. Incompressibility arises from the Pauli Exclusion Principle applied to the densely packed electrons, preventing them from crowding closer together. A degenerate core is not bound by the 10% mass limit discussed above; it can support the entire mass of a 1.3 $M_{\odot}$ star while still remaining degenerate and isothermal because of its high thermal conductivity. Now the degenerate core of such a star is not absolutely incompressit only relatively so, as is a metal compared with a gas. Such a core can contract gravitationally and become hotter as a consequence when more matter (He) is added to it from an H-burning shell. It remains isothermal but grows hotter as the shell burning continues until eventually the 3α ignition point is reached.

Once He-burning begins, the core temperature rises very rapidly. The heat generated by these reactions does not have time to escape so the temperature also rises very quickly, causing the 3α process to go even faster. The rate of core heating consequently accelerates until the pressure due to thermal motions of the nuclei exceeds that due to the degenerate electrons, and the degeneracy is suddenly lifted. When this happens the core "goes normal" again and becomes a conventional hot gas once more. It is as if the "metallic" core were suddenly vaporized. This thermal runaway with removal of degeneracy is termed a "helium flash", and it results in a sudden reorganization of the star's interior structure that produces a corresponding rapid increase in luminosity and radius of the star. Such a sudden, extensive readjustment of the structure of a star ought to be observable, but two factors act to make such observations difficult. Most of the heat of the expanding core at the helium flash is absorbed in heating the outer layers of the star beyond the core where most of the star's mass resides. Thus the helium flash is heavily blanketed by the overlying portions of the star, and the rise in surface luminosity is not very large, certainly far less than the transformation of the core might suggest. The second point that affects observability is that the flash is short-lived, and at any one short period of time, say the past century during which photometric observations have been made, the chances of seeing such an event are rare. While the helium flash peak itself may not ever have been observed, the slower phases leading up to and beyond the flash appear fairly clearly in the H-R diagrams of star clusters. There is one more selective effect: only stars in a narrow mass range between $\approx 1 - 1.3\ M_{\odot}$ appear capable of undergoing a helium flash. Stars more massive than $\approx 1.3\ M_{\odot}$ don't develop degenerate cores; those less massive than $\approx 1\ M_{\odot}$ cannot develop high enough central temperatures to ignite the 3α process and therefore have no central nuclear energy source to produce a helium flash.

Returning to consideration of evolutionary tracks in Figure 5, the path between Points 3 and 6 represents a phase in which the core has ceased to burn H, having run out of fuel, and the energy of the star is supplied mainly by a hydrogen burning shell. This shell eats its way outward in mass but doesn't actually move outward radially. It is a kind of incinerator into which gravity keeps pushing more fuel from outer regions of the star. In fact, as time goes on, the H-burning shell actually contracts in radius making the inert core plus shell more and more compact while the envelope grows radially. The star with an enlarged envelope has a much larger surface area and therefore doesn't require as high a surface temperature to radiate away its energy, so the H-R track moves to the right to lower temperature as the star grows in size.

Stars less massive then $9\ M_{\odot}$ develop deep convective envelopes at Point 5 and behave very much as pre-main sequence stars do as they contract along the "Hayashi track", a phase also characterized by extended convection. The red giant star that is deeply convective evolves along a Hayashi track, but in reverse from Point 5 to Point 6. At Point 6 helium burning begins if the star is massive enough ($\gtrsim 1\ M_{\odot}$) and a helium flash may follow which would add a small nearly vertical spike at Point 6.

Stars less massive than ≈ .5 $M_\odot$ cannot become hot enough to ignite the triple alpha He-burning process, and their cores remain degenerate until eventually the hydrogen burning shell runs out of fuel and extinguishes. Thereafter the star has only limited reserves of gravitational energy, and when these are soon exhausted, the star simply cools slowly as a He white dwarf. In the process the star contracts to a very small radius of the order of a few Earth radii. We have seen the trend toward compaction of the core in red giants in the manner in which a hydrogen burning shell is fed from the outside by material pulled inward by gravity, while at the same time the shell actually contracts in radius. This characteristic trend in the giant phase toward concentration of most of the mass of the star into a dense, compact core surrounded by a distended, tenuous atmosphere reaches extremes in white dwarfs. In these stars the cores are all that remain except for just a trace of atmosphere bound closely to the surface by the strong surface gravity of so much mass in such a compact ball of matter. Virtually the whole star is degenerate, therefore more or less isothermal and rigid - a massive, very hot, very small cinder of a star left to radiate away its thermal energy until it grows cold.

The cooling time for a white dwarf (W.D.) may be calculated simply. A particle (nucleus, ion, electron, atom) in a star at temperature T has thermal energy $\frac{3}{2}$ kT (k is Boltzmann's constant), and a star of mass M_{gms} consisting of pure He, fully ionized, has total thermal energy aTM where a is a constant that includes k and the mean mass per particle of stellar material ($He^{++} + 2e^-$). The luminosity L is furnished by radiative loss of the thermal energy so

$$L = \frac{d}{dt}(aTM) = aM\frac{dT}{dt}$$

The theory of the structure of white dwarfs shows that $L \simeq KT^{3.5}$ where K is another constant. Eliminating L between these two equations gives

$$\frac{dT}{dt} = \frac{K}{aM} T^{3.5}$$

This has a simple solution, $\tau = AMT^{-2.5} = BMT/L$ and A, B are constants. The time, τ, calculated in this way is a cooling time, roughly the time it takes the star to lose a significant fraction of its initial thermal energy. The table below gives cooling times for He white dwarfs.

Table III. Cooling Times of He White Dwarfs

$M/M_\odot$	$L/L_\odot$	τ (years)
.5	10^{-2}	9×10^9
.5	10^{-3}	4.8×10^{10}
.5	10^{-4}	2.4×10^{11}
.25	10^{-2}	5.5×10^9
.25	10^{-3}	2.9×10^{10}
.25	10^{-4}	1.5×10^{11}

All of these except the two with $L/L_\odot = 10^{-2}$ have cooling times that exceed the age of the Universe so we can expect the W.D. phase of a star's lifetime to be comparable with its M.S. phase or much longer.

Observed T and L values of white dwarfs place them well below the M.S. and somewhat to the left at high temperatures. As they cool, they move along H-R tracks of nearly constant radius due to the relative incompressibility of degenerate matter. Such tracks roughly parallel the main sequence and lie below it by factors 10^{-2} - 10^{-4} in luminosity.

Returning again to the He-core burning stage of the giants massive enough to have ignited the triple alpha phase, once this new energy source comes into operation the star settles down again into a fairly normal structure rather like an UMS star. In fact the evolutionary track from Points 6 - 8 is very much like a replay of the initial approach of the pre-MS star to the MS along a Hayashi track followed by its long stable residence on the Main Sequence. The inner structure of the star at this phase is of course more complex than a MS star, for it has not only a turbulent convective He-burning core but also a H-burning shell surrounding the core. When He begins to fuse in the core, the latter expands and core expansion is accompanied by a contraction of the distended red giant envelope. The resultant slow rise in surface temperature moves the star leftward in Figure 5 to Point 8, a process requiring a few million years for stars in the 5 - 15 $M_{\odot}$ mass range. At Point 8 He exhaustion occurs in the core. This happens all at once throughout the convective core much as H exhaustion in the UMS stars occurs. As the star evolved to the right in Figure 5 in the latter case, so it does in the present stage when He exhaustion takes place in the core. The path on the H-R diagram is from Point 8 to Point 9 in Figure 5. An inert carbon core with appreciable amounts of oxygen is left after He-burning. This inert core will generally be surrounded by a He-burning shell and, farther out, a H-burning shell. If the star is massive enough, it may attain central temperatures high enough to ignite carbon fusion to yield a variety of heavier products. If the core is degenerate when this happens a "carbon flash" will occur similar in some respects to the He flash.

As stars between about 5 and 10 solar masses swing back and forth on the H-R diagram, in their various giant phases, they cross a region of instabili marked by dashed lines in Figure 5 and designated as a region of classical cepheids. These are radially pulsating stars whose luminosities vary periodically with periods that are related to the mean luminosity (the higher the luminosity, the longer the period). These stars have played an important role in observational astronomy as distance indicators. They are bright and readily recognizable. Since their measured periods of brightness variation serve to determine their luminosities, measurement of their apparent magnitudes, or brightness as seen from the Earth, serves to determine their distances.

The scenario of further evolution of massive stars can be inferred from what has been said above. If a star with a carbon core is not massive enough to ignite carbon burning, it may eventually become a pure carbon white dwarf consisting almost entirely of carbon nuclei and degenerate electrons much like the helium W.D. discussed above, or it may have a more spectacular fate in store as discussed below.

If C-burning is ignited, the star then makes another left-right, looping excursion on the H-R diagram similar to those following earlier core ignitions. Less energy is released from these carbon and subsequent burnings so they take less time; evolution speeds up rapidly, and each new burning phase leads subsequently to a new shell burning phase. The structure of the star becomes more complex as layer develops upon layer. But there are two possible ends to this sequence.

As evolution proceeds the central temperature continues to rise, and when it reaches approximately a billion degrees, the star begins to lose energy in significant amounts by the emission of neutrinos. Neutrinos are produced in all the earlier nuclear fusion processes, of course, and carry away some energy, but energy loss by neutrino emission amounts to a few percent which is not enough to upset the star's structure or to affect its energy balance appreciably. However, above a billion degrees several new processes become possible that are copious producers of neutrinos, and these processes increase very rapidly with temperature. The significance of neutrino production is that neutrinos, once produced, hardly interact with matter at all. In fact a neutrino may traverse ≈ 3000 light years of lead before interacting with a lead nucleus! Consequently they are able to leave the star without being stopped and are thus able to transport energy directly out of the central regions of the star at the speed of light with no hindrance. A normal main sequence star suffers an energy leak of some 2 - 6% by neutrino emission, but a massive star in the late stages of its evolution produces neutrinos at much higher rates than a main sequence star as a result of processes other than normal nuclear reactions that take place at very elevated temperatures and densities. The enhanced neutrino emission can quickly remove large amounts of energy from the star triggering a collapse. This neutrino loss has been referred to as a sudden "refrigeration of the core" that drops the pressure in the core which sustains all the overlying stellar material. With its foundations suddenly removed, the outer matter collapses inward in a stellar implosion. The refrigerated core rapidly shrinks to a degenerate ball upon which the infalling stellar matter impinges, bouncing back in a cataclysmic shock wave probably enhanced by nuclear burning of infalling fuel (H, He, etc.) suddenly thrust into the billion degree interior of the star. The resulting explosion gives rise to a supernova whose peak luminosity ($\approx 10^{10}\ L_{\odot}$) may well exceed that of its entire parent galaxy. The explosion divests the core of most of the outer region of the star, usually several solar masses of material, and produces an active, irregularly expanding nebula typified by the Crab Nebula.

The extremely compact, degenerate core that remains is a neutron star, probably rotating rapidly and observable as a pulsar. This neutron star represents a different order of degeneracy than appears in the white dwarf. In the latter the electrons were so tightly packed that quantum effects associated with the Pauli exclusion principle began to restrict the way in which they could be packed. In the neutron star few electrons remain, nearly all having combined with protons to form neutrons. But now the latter are so densely packed that their packing becomes restricted by the same Pauli principle applied to the neutrons rather than the electrons. The net result is similar, however. The degenerate neutron matter behaves much as the degenerate electrons in a white dwarf in that it is virtually incompressible and has high thermal conductivity. It also has some other interesting quantum properties such as superfluidity and superconductivity according to some theorists.

Degenerate electrons are incapable of supporting greater gravitational pressure than that produced by about 1.4 $M_{\odot}$ so a degenerate core more massive than this cannot be a stable white dwarf and is destined to collapse to the next stage of degeneracy at the neutron star stage. A star whose initial mass lies between 1.4 and approximately 2 solar masses is believed to form a WD eventually by shedding all its mass in excess of 1.4 $M_{\odot}$. It does this late in its red giant stages mainly via radiation pressure blowing away its distended outer atmosphere in a kind of enhanced "solar (stellar) wind". Red giants are believed capable of losing as much as 10^{-6} $M_{\odot}$ per year in this way, and the giant phase lasts long enough for the star to shed more than 1 $M_{\odot}$ to reduce itself below the 1.4 $M_{\odot}$ limit for white dwarfs. This suggests that the usual fate of stars initially less massive than $\approx 2\ M_{\odot}$ is to become white dwarfs, and observational counts of WD's tend to confirm this view.

The process of mass shedding mentioned above results in ejection of stellar matter with an outward radial velocity of several tens of kilometers per second, and the ejected matter forms a spherical nebulosity around the central star. It is very tenuous material and is optically excited by ultra-violet from the hot central star. As a result the ejected matter appears to a distant observer not as a spherical ball of glowing gas but rather as a thick, faintly glowing, ring, resembling a smoke ring centered on the parent star. These planetary nebulae, as they are called, are beautiful to behold and are sufficiently numerous to support the view that most stars somewhat heavier than the sun up to $\approx 2\ M_{\odot}$ form them late in their evolutionary life.

The fates of stars $\gtrsim 2\ M_{\odot}$ has been described as explosive, ending as supernovae that leave behind neutron stars and lose much mass in the explosion. Observations of the rates of occurrence of supernovas have a large uncertainty, but rates at the upper end of the observed range are consistent with the view that all single stars $\gtrsim 2\ M_{\odot}$ become supernovas.

Supernova explosions play an important role in the evolution of galaxies, for they provide a means whereby the interstellar medium of gas and dust becomes gradually enriched in heavier elements than hydrogen and helium that we believed to be the initial material from which the galaxies and their first stars were formed. There is no doubt that the envelope ejected in a supernova explosion will be enriched in the normal products of nuclear fusion: He, C, O, N, Ne, F, Mg, Si, Fe and some intermediate elements. There are also nuclear process that take place in the short time during which the infalling gas is heated to billions of degrees. These processes build up other elements lighter than Fe besides those listed and also build up elements heavier than iron all the way to uranium and a little beyond. In fact, SN explosions may well be the only way in which an element like gold can be formed. It is sobering to reflect that one's ring or the fillings in one's teeth were born in the heart of one of nature's most awesome events! Stars subsequently formed from this material enriched with heavier elements evolve somewhat differently in certain details than their earlier generation counterparts, and clues as to the ages and evolution of galaxies are contained in the relative numbers of metal-rich and metal-poor stars.

In discussing neutron stars a hierarchy of degeneracies was mentioned: first electron degeneracy stabilizing the white dwarfs, then at greater densities there follows neutron degeneracy stabilizing neutron stars. Is there a limit to the mass that neutron degeneracy can support as there was for electron degeneracy? The answer is a definite yes, but the exact limit of the mass is not known as precisely for the neutron star as for the white dwarf because the properties of neutron matter are not well understood. The limit probably lies somewhere between 2 and 3 $M_{\odot}$, and a neutron star more massive than this appears to have no further brakes to its collapse. It shrinks until its gravitational pull will not even allow light to escape from its surface, and it becomes a "black hole", effectively cut off from our Universe by all save its gravitational pull on other bodies. This must surely be the most extreme fate conceivable for a star.

So far we have considered only the evolution of single stars unaffected by any neighbor. Stellar evolution takes on a new degree of complexity when we consider evolution of pairs of stars bound together by gravity into close binaries. Such binaries constitute over half of all the stars.

Consider a close binary system of two stars of somewhat different masses on the main sequence. The more massive star evolves into the red giant stage first, and in the process its swollen atmosphere comes under the gravitational influence of the other member of the binary. When this happens, a sizeable fraction, and in some cases all, of the atmosphere of the giant may be pulled off by the less massive star. This material forms an accretion disk about the latter star and leaves behind the hot core of the former giant, now in the form of a helium white dwarf.

The star that has accreted the former giant's atmosphere is now more massive than before, and its evolution to the red giant phase is considerably speeded up. When it expands to the red giant configuration, its envelope comes under the gravitational influence of the WD companion and may be captured by the latter, reversing the original exchange of matter. When this happens, the star that has recaptured its envelope may be able to proceed on its evolutionary course, and if it is massive enough to undergo a second stage of red giant expansion (when He burning ceases in the core), another transfer of envelope may take place.

The extent of envelope transfer and the number of times it can take place depends upon the masses of the stars of the binary system and upon their operation. With these additional variables the variety of possible evolutionary scenarios becomes quite large indeed. But there are still more possibilities to complicate the picture. Suppose one of the stars becomes a supernova which can eject a considerable amount of matter from the system. If the mass ejected is less than half the total mass of the system, the binary stars will remain together as a binary, albeit with quite eccentric orbits. If the eject mass exceeds the above limit, the binary system will be disrupted and each star will go its way as a relatively high velocity single star.

Under proper conditions accretion of mass by one member of the binary can lead to formation of a black hole. At the present time the most likely candidate for an observed black hole, is an x-ray emitting binary, Cygnus X. Such a star must originally have been too massive to form a degenerate core, and in a supernova explosion it must have formed a core too massive to stop at the neutron star stage. Such stars were probably initially above 3 $M_{\odot}$, though this lower limit is rather uncertain at our present state of knowledge.

In the process of accreting its partner's envelope a star pulls the acquired gas rapidly onto its hot surface, and the process is often energetic enough to produce copious x-ray emission. Thus many binary stars should become strong x-ray emitting sources during accretion phases of their evolution. If one member has evolved into a black hole, it should be a particularly vigorous x-ray emitter since the gas accreted by the black hole is accelerated to near the velocity of light as it approaches the black hole "surface", and collisions between such energetic gas atoms produce not only hard x-rays but gamma rays as well.

Accretion under less dramatic circumstances can also produce rather spectacular results in the form of nova outbursts. Novae are believed to occur only in close binary stars, and current opinion ascribes these impressive mass ejections of novae to explosive burning of hydrogen that is rapidly accreted onto a very hot core such as the He white-dwarf left after initial envelope transfer. Nova explosions are far less energetic than supernova events, and they eject relatively small amounts of matter, of the order of 0.1 $M_{\odot}$ or less. This mass ejection can be accounted for by accretion theories, but it should be emphasized that present theoretical understanding of details of binary evolution is in an early state, and much work remains before the complications of the life histories of these stars are unravelled satisfactorily.

STELLAR STRUCTURE AND ITS DETERMINATION

I. Introduction

Man's ability to deduce the structure and conditions of temperature and pressure inside a star that can only be observed from a great distance must surely rank as one of the crowning achievements of twentieth century science. Present knowledge of stellar structure and its change with time rests upon a large base of astronomical observation and makes use of knowledge from many branches of physics.

The term stellar structure properly includes both the study of the interior of a star and its outer layers or atmosphere, and since these two topics employ different methods of study, there has been a divison of the subject into two special fields, interiors and atmospheres. Most of what we know about a star is learned from the light that it emits, and this light is generated in the outer reaches of the star, in its atmosphere. It is therefore of vital importance to understand the processes that go on in the stellar atmosphere in order to be able to relate the observed parameters such as luminosity, effective temperature, and chemical composition to these variables deeper inside the star. The stellar atmosphere is a boundary layer between the interior and the space into which the star radiates such information as we are priveleged to receive, and the properties of this boundary region must be understood if we are to be able to interpret observations in terms of interior structure.

By good fortune, it turns out that we don't need very detailed knowledge about the properties of the stellar atmosphere to deduce rather good models of the internal structures of most stars. By most stars we mean those on or near the main sequence, the white dwarfs, and red giants in the early stages of their evolution away from the main sequence. The distended atmospheres of older red giants and special features of the atmospheres of some other types of stars in late stages of evolution play important roles in determining the deeper structures of these stars. Details of the atmospheres of such stars must be taken into account if their structures are to be reliably deduced. We shall confine ourselves to the study of the internal structures of normal, main sequence stars for the most part, and for these the effect of the atmosphere can be simulated rather simply. We shall also be interested to see how the interior of a star changes with time as more and more of its hydrogen fuel supply becomes converted by nuclear fusion into helium and heavier elements. These evolutionary effects of exhaustion of the nuclear fuel can be followed as the star leaves the main sequence. Results of extensive computer-aided calculations of the changing structure of the star give a picture of the various scenarios a star may be expected to follow in its later life depending upon its initial mass and composition. This constitutes the basis of stellar evolution. The main topic we will pursue, however, is an exposition of the general method by which details of the internal structure of a main sequence star can be found. The physical basis for these calculations turn out to be disarmingly simple, and once the basic stellar structure equations have been derived, mainly through straightforward physical reasoning, they can be used to calculate some simple stellar models.

A stellar model is here taken to mean the radial dependence of several pertinent variables, usually temperature, pressure, density, composition, and the net energy flow outward through a spherical surface at radius r within the star. The mass contained within radius r is also generally determined as a function of r. These variables presented either in tabular or graphical form constitute a stellar model.

The object of stellar interior calculations is to devise a model whose total mass, M, luminosity, L, and surface temperature, T, correspond with a real observed star. It might be suspected that many different models could give the same M, L, and T_s. If this were the case, how could the correct model be identified, if indeed there were a "correct" model? Under these circumstances it would be hard to do any meaningful stellar interior astrophysics. The field of model calculations would lack the essential requisite of any good science, namely unique contact with observation. Fortunately for the development of our knowledge of stellar interiors and evolution, it turns out that with a few rare pathological exceptions, the structure of a star is uniquely determined by its total mass and chemical composition. Thus, we are assured that if we can calculate a model that gives the same observable properties as a real star, e.g., M, L, T_s, and composition, μ, this model will be unique and therefore the correct structure of the corresponding real star. This uniqueness theorem is due to Vogt and Russell and is a consequence of the mathematical structure of the equations that determine the equilibrium internal structure of a star. A more detailed discussion of the Vogt-Russell theorem and its limitations can be found in Cox and Guili.(1)

II. Observational Base

Let us consider briefly the astronomical observations on stars that make up the data base upon which the theory of stellar structure rests and against which the results of such theoretical calculations must be tested. These observations can be grouped into several classes:

a. Mass

Stellar masses are determined from measurements of the dynamics of binary star systems. The measurements include a variety of of different types of normal stars, but the number of precise measurements is not very large, numbering perhaps a few hundred, so this body of observational data is rather limited.

b. Luminosity

The inverse square law applied to the apparent brightness of a star observed photometrically on Earth gives its luminosity, L, provided its distance is known and provided that its total radiant energy is measured by the astronomer. Usually, the total radiation cannot be measured, but only the radiation falling into a limited wavelength range. In this case the total radiation can usually be inferred from the spectral distribution of the radiation actually measured in the restricted region of the spectrum. Except for relatively nearby stars whose distance can be measured accurately by triangulation (some 6000 stars), stellar distances must be determined by various indirect methods subject to more or less error. In the matter of deducing the total radiation of a star over the whole spectrum from photometric measurements made in the visible part of the spectrum, there is considerable room for error especially for very hot stars that radiate much ultraviolet light and for cool, red stars that emit much infrared radiation.

c. Temperature

The surface temperature of a star (actually the temperature of the relatively thin shell, the photosphere, where the visible radiation of the star originates) can be determined by observing the color temperature, or by deducing the temperature from relative intensities of several spectral lines, or from a knowledge of the star's radius (rarely known from independent measurement, with the notable exception of the Sun). The surface temperature is most often found by measuring the color temperature, i.e., the temperature of a blackbody that would have the same color (spectral distribution) in the visible region of the spectrum as the star under observation. The temperature in the interior of a star below the photosphere is inaccessible to observation, of course,

d. Radius

Except for the Sun whose diameter can be measured by triangulation and a small number of nearby stars, mostly large red giants, whose diameters have been measured interferometrically, and a number of eclipsing binaries whose light curves give the diameters of the components, we have no other direct observational knowledge of stellar radii. However, the luminosity, surface temperature, and radius are related by the Stefan-Boltzmann law, $L = 4\pi \sigma R^2 T^4$. This assumes that the star radiates as a blackbody which is a reasonably good assumption for the majority of main sequence stars. Departures from blackbody radiation can be taken into account by defining an effective blackbody temperature, T_{eff}, to be used in place of T_s. In either case the stellar radius is related to luminosity and temperature by a Stefan-Boltzmann type of equation.

e. Chemical Composition

The Vogt-Russell theorem identifies composition and mass as the main factors that determine the structure of a star, so the role of composition is certainly of great importance. It is unfortunate that only the surface composition can be determined by observation. Stellar spectroscopy is used to determine the surface composition and it is applied in much the same way as it is used in the laboratory to identify elements and simple molecules from the wavelengths of their characteristic spectral lines or bands. These are seen in stellar spectra mainly in absorption (or occasionally as emission lines in some very hot stars). The relative strengths of the absorption lines in a stellar spectrum give quantitative information about the concentrations of various elements in the outer layers of a star, but one must rely upon educated guesswork to obtain from this surface composition an idea of the internal composition. The two are related in different ways in different types of stars, depending to a large degree upon the amount of mixing of the surface with underlying material throug[h] the action of turbulence and convection. In some stars this mixing is important, and in others it is not. It is therefore a matter for expert judgment to relate observed chemical abundances at the surface of a given star with the interior composition needed to determine unambiguously the star's internal structure. The simplest assumption is that the stellar material is well mixed in the early stages of formation of the star and that the star begins life on the main sequence with a homogeneous composition, uniform throughout and the same as that observed in later life in its surface layers. This means simply that the observed surface composition is assumed to be the same as that of the interior of the star when it begins life on the main sequence and starts converting hydrogen into helium near its center.

As time goes on, the internal composition will change, of course, because fusion converts lighter elements into heavier ones, but these processes proceed at definite and predictable rates. If we know the initial composition, we can keep track of how it changes with time at every point in the star. We rely upon the absence of nuclear processes in the outer layers of the star to preserve the primordial composition there where the astronomer can measure it spectroscopically. This scenario does not hold universally, but it appears well justified in the majority of stars and serves as a useful working hypothesis. On the basis of spectral analysis of stellar light we expect the mass concentrations of most stars to consist primarily of hydrogen ($\approx$70%) and helium ($\approx$25-30%) with relatively little ($\leq$3%) of heavier elements, chiefly neon, oxygen, nitrogen, with still smaller amounts of carbon, magnesium, silicon, sulfur, and iron ($\approx$0.1% each). Other elements than these are usually present only in trace amounts in the outer layers of the star, but deep inside where nuclear transformations take place, the heavy elements may build up appreciable concentrations in time.

Besides the observed features listed above, there are others of a generally different character whose bearing upon the subject of stellar structure is either less direct or less well understood at present. Among these are the surface magnetic fields and rotational rates of stars. Most stars have small magnetic fields similar to the field of the Sun and amounting to less than about 100 gauss at the poles. Some stars, however, have strong fields in the kilogauss range, while fields at the surface of white dwarfs are thought to range from $\approx 10^5$ to as high as 10^8 gauss.

Stellar magnetic fields are measured spectroscopically by observing the splitting of spectral lines through the Zeeman effect. Technically the measurements are very difficult except in the stars with very strong fields because in most weak-field stars the spectral lines are not very sharp. They are broadened by effects of collisions and thermal velocities, by turbulent motions of the photosphere, and by Doppler shifts originating from stellar rotation. The Zeeman splitting is generally only a minor contribution to overall line shape in such stars and is obtainable with considerable uncertainty only after all the other broadening effects have been sorted out and quantitatively accounted for. It is fortunate for progress in the field of stellar interiors that the magnetic field of a star seems to play only a small role in determining the internal structure of a normal star. Present knowledge about the origin of the field in a star like the Sun indicates that the field originates from convective motions of ionized material fairly far out near the photosphere, at most involving less than the outer 20% of the Sun's radius. In this region the interaction of the field with the turbulent gas should affect the scale and detailed dynamics of the turbulence. Considering how little we know at present about turbulence in general, it is safe to say that magnetic interactions do not seem to play a major role in determining the structure of the outer convective layers of a solar type star. Whether this holds true for stars with much stronger fields is open to debate.

Rotation has been grouped with magnetic fields for a good reason. The two appear to be intimately connected, and in the case of the Sun, Parker [2] has shown in detail how the solar field arises out of dynamo effects driven in part by differential rotation of the various layers of the Sun. It is also strongly suggestive that the strong-field stars are all rapid rotators as well. In fact, observation reveals that the younger, more massive hot blue stars (those "earlier" than spectral type A, having surface temperatures $>8000^\circ K$ and masses greater than about 1.5 solar masses) are the rapid rotators and include all the known strong-field stars. Rapid rotation, besides giving rise to strong fields which should influence convective, turbulent regions of the star, also makes the star non-spherical. Rapid rotators are oblate because of centrifugal forces and can bulge very considerably at the equator. This means that the spherical symmetry that is so helpful in reducing complexity of the general equilibrium structure equations no longer holds, and the structure calculations become much more difficult. Even so, in the cases in which rotational distortion has been taken into account, the structures do not differ greatly from the non-rotating stars. Rapid rotation might be expected to influence the rate of mass loss from stars via the "stellar wind", a counterpart of the solar wind of matter ejected from the Sun. Mass loss is particularly pronounced in the hotter, more massive stars where the radiation pressure is so great due to the high radiation density at the surface of the star that significant amounts of matter are blown away. Rapid rotation should enhance this mass loss from the equatorial regions of the star where centrifugal force aids radiation pressure.

Observationally, stellar rotation is measured by the symmetric broadening of spectral lines due to the Doppler shift. Every surface element of the star is either approaching or receeding from the observer at a velocity which depends upon its position on the stellar surface, upon the rotational velocity of the star, and upon the angle between the observer and the axis of rotation of the star. It is possible, by separating this rotational part of the spectral line shapes from other sources of line broadening, to arrive at a star's rotational velocity. The data actually give $\Omega \sin \theta$ where Ω is the rotational velocity and θ is the inclination of the axis of rotation. Statistical studies of large groups of stars of similar type serve to separate Ω from θ since the latter may be assumed distributed randomly over all angles. Mass loss can sometimes be measured directly, the best case being planetary nebulae that have ejected mass in the form of a visible shell whose extent can be seen and whose density can be inferred from spectroscopic data on the shell radiation. Mass loss from certain other types of stars such as the very luminous, young blue O-stars and the older, post main-sequence K-giants can often be deduced from spectral measurements of gas ejection velocities, again making use of the Doppler effect.

There remains to be mentioned one unique observational means for "seeing" into the center of a star, neutrino astronomy. Neutrinos are ejected in the nuclear fusion processes that provide a star's energy, and the rate and energy of neutrino emission are calculable. Since neutrinos are hardly stopped at all by matter, they travel outwards from the center of the star quite unimpeded and off into space. The Sun is a close enough source, and the only source close enough, to generate a measureable neutrino flux on Earth. Measurements of this flux are very difficult because of its small interaction with matter, but Davis and his coworkers [3] have succeeded in establishing the fact that the solar neutrino flux is at least a factor 4 less than that predicted by current models of the Sun, and therein lies an important current problem in astrophysics to be solved.

Finally, all of the observational information implicitly contained in the H-R diagram constitutes a primary pool of data against which to test theoretical stellar models and their evolution with time. The H-R diagram is really quite a restrictive observational constraint. Any valid stellar model must fit somewhere on this diagram in regions where real stars are found, and it must conform to the other known properties of real stars in that part of the diagram. It must, for example, also fit the observed mass-luminosity relation for real stars, another restriction on possible models.

III. Equations of Stellar Structure

As in most other branches of physical science the astrophysical determination of stellar structure and evolution proceeds by a regular interplay between theory and observation. Starting with initial, reasonable assumptions about chemical composition, a model star is calculated, and its mass, luminosity, surface temperature and any other observable properties are calculated and checked against observed stars to see whether the model star matches any real star. If not, the assumptions are changed, a new model is calculated, and it is again compared with the gamut of real stars. When the model matches a real star, the Vogt-Russell theorem assures its correctness provided conditions for the applicability of the theorem are met. These are quite general, and the theorem almost always holds.

We now proceed to develop the principle equations that govern the interior structure of a spherically symmetric star in equilibrium. By equilibrium we mean that structural changes may take place, but only slowly on a time scale of at least the order of 10^7 years which is the time it takes a star like the Sun to readjust itself by the action of gravity to small structural disturbances. This time scale is called the Helmholtz-Kelvin contraction time, T_{HK}. It is one of several important stellar time scales and the one which measures the relaxation time of a star for departures from energy balance (i.e., thermal equilibrium). It is in order of magnitude,

$$T_{HK} \approx \frac{E_{th}}{L} \tag{1}$$

where E_{th} is the total internal thermal energy of the star. If the star has no strong magnetic fields, and if its internal energy is made up mainly of non-relativistic, monatomic particles, E_{th} can be evaluated in terms of M and R for the star. The resulting expression is approximately

$$T_{HR} \approx 2 \times 10^7 \left(\frac{M}{M_\odot}\right)^2 \left(\frac{L_\odot}{L}\right) \left(\frac{R_\odot}{R}\right) \text{ years} \tag{2}$$

where $M_\odot$, $L_\odot$, $R_\odot$ are respectively the solar mass, luminosity, and radius. The significance of T_{HK} is that the radius of a star adjusts to non-equilibrium conditions at the rate

$$R(t) = R(o)\, e^{-t/T_{HK}} \tag{3}$$

Another time scale of interest to us is the nuclear time scale, T_N, that measures the time it takes for the properties of a star to change significantly as a result of nuclear reactions in the interior. This time, too, will have dimensions of an energy divided by the luminosity, in this case E_N/L where E_N is the available nuclear energy. For a star that derives its energy from conversion of H^1 to He^4, a process with an efficiency of 0.7% of the total mass-energy Mc^2, one would expect E_N to be proportional to $.007\ Mc^2$. The proportionality factor is about 0.1 because it has been found that normal main-sequence stars begin to change into red giants when about 10% of their mass has been converted into helium. Thus we find

$$T_N \simeq 7 \times 10^{-4} \frac{Mc^2}{L} \simeq 10^{10} \left(\frac{M}{M_\odot}\right)\left(\frac{L_\odot}{L}\right) \text{ years} \qquad (4)$$

Thus T_N and T_{HK} are of the order of 10^{10} years and 2×10^7 years respectively for a star like the Sun, though for very massive stars both time scales become much shorter. With the help of the observed mass-luminosity relation, L can be eliminated from these time scales. The M-L relation is usually given in the form $L \propto M^n$ where n depends somewhat upon mass and ranges from about 2.75 for low mass main-sequence stars to about 4 for stars more massive than about 0.6 $M_\odot$. Observational data on stellar radii also lead to the empirical relation $R \propto M^a$ where a also varies but is approximately 1 for main-sequence (MS) stars more massive than $\simeq 0.6\ M_\odot$. With these M-L, M-R relations the time scales can be expressed in terms of M alone:

$$T_{HK} \simeq 2 \times 10^7 \left(\frac{M_\odot}{M}\right)^3 \text{ years} \qquad (5)$$

$$T_N \simeq 10^{10} \left(\frac{M_\odot}{M}\right)^3 \text{ years} .$$

or $\frac{T_{HK}}{T_N} \approx 10^{-3}$ for all MS stars more massive than $\approx 0.6\ M_\odot$.

A normal star can therefore adjust its structure to achieve energy balance and hydrostatic equilibrium on a time scale that is very short compared with the scale of major evolutionary changes, T_N. This conclusion establishes the validity of treating normal stars as "equilibrium" objects bound together by gravity on one hand and prevented from collapsing on the other hand by the outward flow of energy from nuclear sources in the interior and by the immense pressures generated in the star by the thermal energy it contains.

This quasi-equilibrium property of the star greatly simplifies calculation of its structure, for it allows us to apply relatively simple static equations to the interior. Consider first the requirement that the star be in hydrostatic equilibrium which means that a given blob of gaseous material in the star is at rest, balanced by the inward pull of gravity on one hand and by the outward buoyant forces of the pressure gradient on the other. This buoyant force is exactly analogous to the familiar buoyant force in any fluid. Archimedes Principal says that this force is equal to the weight of the displaced fluid which is simply the mass of the fluid times the gravitation acceleration at its rest position.

Suppose we take as our sample of gas in equilibrium a thin spherical shell in the star. Let it have radius r and thickness dr as shown in Figure 1:

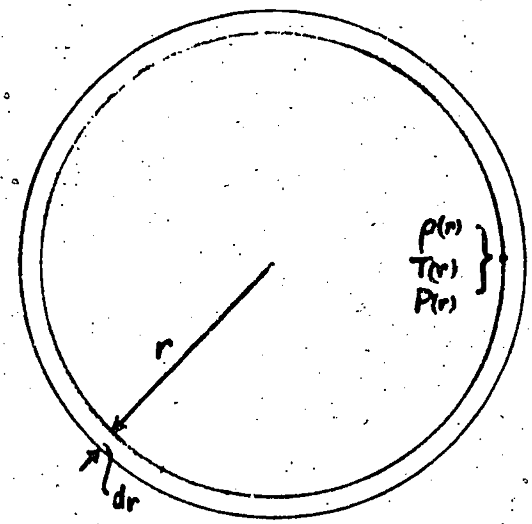


Figure 1. Thin spherical shell inside a star for which conditions of equilibrium are applied.

Let the material within this shell have density, $\rho(r)$, temperature, $T(r)$, and pressure, $P(r)$. The mass, $dM(r)$, of material in the shell is given simply by

$$dM(r) = 4\pi\, r^2\, \rho(r)\, dr \qquad (6)$$

and the weight (i.e., gravitational force acting on the shell) is just $g(r)dM$ where $g(r)$ is the gravitational acceleration acting at radius r.

In a spherically symmetric distribution of matter the gravitational pull at any point r from the center is as if all the matter inside a sphere of radius r were concentrated at the center. If G is the gravitational constant (6.67×10^{-8} dyne cm^2 gm^{-2}), it follows that the gravitational force on (dM) at r is, making use of (6),

$$F_g = \frac{GM(r)}{r^2}\, dM(r) = 4\pi\, GM(r)\, \rho(r)\, dr \qquad (7)$$

where M(r) is all of the mass lying within radius r; that is,

$$M(r) = \int_0^r dM(r) = \int_0^r 4\pi\, r^2\, \rho(r)\, dr \qquad (8)$$

The buoyant force on dM(r) is supplied by the difference in pressure between the inside surface of the shell where the pressure is P(r) and the outside surface of the shell where it is $P(r + dr) = P(r) + \frac{dP}{dr} dr$. Pressure, being force per unit area, exerts a total outward force of $4\pi r^2 P(r)$ on the inside surface of the shell and a net force, due to the pressure gradient, $dP(r)$, of

$$F_p = 4\pi r^2 \, dP(r) \tag{9}$$

It is this net outward force on the shell that is balanced by the inward pull of gravity given by (7). At equilibrium, therefore (7) and (9) must be equal, so

$$dP(r) = -G \frac{M(r) \rho(r)}{r^2} dr \tag{10}$$

The minus sign has been added to signify that the pressure decreases as r increases. It must eventually go to zero at the "surface" of the star.

This equation (10) together with (6) or (8) defining M(r) make up two of the four principal equations that determine a normal stellar structure.

The third structure equation deals with the energy generation in the star and its connection with the total outward energy flow, L(r). This quantity will be called the luminosity at r by analogy with the total energy radiated by the star which we have termed its total luminosity, L.

Let $\varepsilon(r)$ be the rate at which each gram of matter in the shell of radius r in Figure 1 is generating energy. Then, since there are $4\pi r^2 \rho \, dr$ grams of matter in the shell, the net contribution of the shell to the luminosity is

$$dL(r) = 4\pi r^2 \rho(r) \, \varepsilon(r) \, dr \tag{11}$$

This is the third principal structure equation and the first so far to introduce explicitly a property of the material of the star, ε.

The remaining structure equation deals with the processes by which energy flows outward in the star and is therefore expected to contain additional properties of the stellar matter. The three principal modes of energy transport are radiation, convection, and conduction. We can rule out the latter as playing only a very minor role in most stars (exceptions will be discussed later) by noting that even at the immense pressures and densities present in the central regions of ordinary stars ($\approx 2 \times 10^{11}$ atmospheres and ≈ 160 gms/cm^2 in the Sun) the temperatures are sufficiently high ($\approx$15 million degrees in the Sun) to maintain matter in very nearly an ideal gas state. Likewise conditions outward from the center also confirm the validity of the ideal gaseous state throughout the star, again with some exceptions to be discussed later. Gases are known to be poor conductors of heat so we expect that energy will be transported through a gaseous star either by radiative transfer (stepwise absorption and reemission of radiation) or by convection (turbulent vertical motions of the gas with hotter gas from below rising to mix and transfer its heat to the cooler overlying gas, the cooler gas sinking in the process to become further heated below). Usually one or the other of these processes will predominate in a given region of a star, though both may be going on within the same star. The predominance of one form of energy transport over the other in a given stellar shell allows each process to be treated separately and greatly simplifies matters.

Radiative transfer will be considered first. As mentioned earlier, this mode of energy transfer involves absorption of radiant energy by stellar matter followed by reemission of the energy as radiation again a short time later. The process is not very efficient if only a single step is considered because very nearly as much energy is reradiated inward as outward from any small region, but the absorption and reemission is rapid, so many steps can take place in unit time, and the slight predominance of outward over inward reemission, a consequence of the spherical geometry and temperature gradient in the star, suffices to make the overall process an effective means for getting energy to the stellar surface.

The radiation need not be visible light; indeed, it is not, as soon as one goes a short distance into the star. There temperatures rise rapidly until most of the radiation is ultra violet light, x-rays, and eventually gamma rays near the center. The primary radiant energy given off in the nuclear fusion reactions is in the form of gamma radiation, and it is only after a large number of absorptions and reemissions that the energy of this radiation is degraded into the range we familiarly refer to as "light". It is sometimes remarked that the center of a star would appear dark to a human observer. This is not true because other radiative processes take place involving charged particle collisions that produce visible light (bremsstrahlung and Compton scattering, for example) but it is true that the central region of a star would appear far less bright than its temperature would suggest.

The luminosity, $L(r)$, gives the total energy flowing outward through a spherical surface of radius r inside the star. The outward energy flux at r is therefore just $L(r)/4\pi r^2$. The temperature gradient dT/dr across a shell of thickness dr must be sufficient to maintain this radiant flux. We can get the required dT/dr by a simple argument.

Consider the radiant energy flux $L(r)/4\pi r^2$ as being propelled outward by radiation pressure, P_{rad}, in much the same way as hydrostatic pressure forces water through a pipe. If the radiation pressure is P_{rad} at r and $(P-dP)_{rad}$ at $r + dr$, then it is the difference, $-dP_{rad}$, that gives rise to the outward flow of radiation, and this differential pressure must be exactly balanced by the resistance met by the radiation along its outward path. The resistance is provided by absorption and reradiation processes in the shell dr.

When an atom absorbs radiation the process conserves momentum, so the momentum of the absorbed photon is given to the absorbing atom. If, before it later reradiates a photon, it collides with a nearby atom, some of the momentum it picked up from the absorption of the photon is transfered to the second atom. This net transfer of momentum from photons to gas atoms constitutes a force acting upon the gas arising out of the radiation pressure. A photon of energy E carries momentum E/c which it transfers to a gas atom upon being absorbed. We expect, therefore, that one factor contributing to the resistance of matter to the flow of radiation would be proportional to $L(r)/4\pi r^2c$ (of the form E/c) since this is the net momentum carried by the radiation flux, $L(r)/4\pi r^2$. A fraction of this momentum flow is stopped by the atoms in dr by absorption.

Let the opacity of the gas to radiation be designated by κ (Greek letter kappa, the conventional notation in the stellar interior literature). This is just the usual mass absorption coefficient in physics. If a beam of light passes through a thickness χ of matter of density ρ, the beam intensity decreases according to the familiar equation

$$I_{(\chi)} = I_{(o)} e^{-\kappa\rho\chi} \quad (12)$$

which is equivalent to saying that $dI/I = -\kappa\rho d\chi$. Applying this definition of κ to our spherical shell dr, a fraction $\kappa\rho$ of momentum flux $L(r)/4\pi r^2c$ is transfered to the gas by absorption in dr, and it follows that for this to be balanced by the differential radiation pressure, we must have

$$-dP_{rad} = \frac{\kappa\rho L(r)}{4\pi r^2 c} dr \quad (13)$$

Electromagnetic theory provides a simple relation between radiation pressure, P_{rad}, and the energy density of the associated radiation field, E_{rad}; this is

$$P_{rad} = \frac{1}{3} E_{rad} = \frac{1}{3} aT^4 \quad (14)$$

where $a = 4\sigma/c$ for radiation in local thermal equilibrium at temperature T (σ is the Stefan-Boltzmann constant). The differential pressure, dP_{rad}, is therefore simply $\frac{4}{3} aT^3 dT$, and (13) may be written

$$\frac{dT}{dr} = -\frac{3\rho\kappa L(r)}{16\pi r^2 acT^3} \quad (15)$$

This is the final interior equation needed if the star's energy transport is entirely by radiation with convection unimportant. Note that both ρ and κ vary with r since they are functions of the variables $T(r)$, $P(r)$, and composition. The opacity is another constitutive parameter specifically referring to stellar matter that has to be introduced.

If the thermal gradient, dT/dr, is large enough, convection will develop because the underlying gas will be hot enough and therefore have low enough density compared with the gas in the upper part of our shell, dr, for the buoyance of the lower material to overcome gravity. This lighter underlying gas will rise through the overlying gas forming turbulent convective "cells" that transport thermal energy. From another point of view, if $(dT/dr)_{radiative}$ is too small (because $\rho\kappa$ is too small) to allow the energy L(r) to escape by radiative transfer, the star will find an alternative way for the energy to get out, and convection is the alternative mode of energy transfer.

It is possible to devise a simple criterion to test whether convection can take place at any given place in the star. This is done by seeing whether a blob of gas of density and pressure $\rho(r)$ and $P(r)$ from the bottom of the shell dr would rise or sink if moved outwards by the shell thickness dr and allowed to expand adiabatically to the density and pressure of the surrounding gas at r + dr. If it sinks under these conditions, convection will not take place and energy will be transferred by radiation. The criterion that results from these considerations is (4)

$$\left| \frac{1}{T} \left(\frac{dT}{dr}\right)_{rad} \right| < \left| \left(1 - \frac{1}{\gamma}\right) \frac{1}{P} \frac{dP}{dr} \right| \text{ for stability} \qquad (16)$$

If the inequality is not obeyed, convective transport will predominate. In this expression γ is the usual ratio of specific heats at constant pressure and constant volume for the stellar gas ($\gamma = c_p/c_v$). This ratio enters into expression (16) because during adiabatic expansion of the test sample of gas moved from r to r + dr no heat is interchanged with the surroundings, and the adiabatic expansion law, $PV^{\gamma} = P\rho^{-\gamma}$ = constant, must hold.

Once convective transport has been established, a different expression than (15) for dT/dr must be applied because the energy transport rate is different. Schwarzschild[4] gives a simple approximate theory for convective energy flow in terms of the excess temperature that an adiabatically expanded volume of gas moved to r + dr has over the temperature of the surrounding gas at r + dr.

If the blob of gas travels on the average a distance ℓ_m, the "mixing length", before becoming mixed with the surrounding gas and losing both its identity and the energy it carries, the final expression for the convective heat flux can be written[4]

$$H_{conv.} = \frac{1}{4}\rho \cdot c_p \left(\frac{GM(r)}{Tr^2}\right)^{\frac{1}{2}} (\Delta\nabla T)^{3/2} \ell_m{}^2 \qquad (17)$$

where $\Delta\nabla T$ signifies the excess temperature of the rising gas over its surroundings ℓ_m above its starting level. This flux may be compared with the radiative flux,

$$H_{rad} = \frac{L(r)}{4\pi r^2} = \frac{-4}{3}\frac{ac}{\kappa\rho} T^3 \frac{dT}{dr} \qquad (18)$$

obtained by rewriting (15). The total flux $H = H_{rad} + H_{conv}$ can be solved for the actual temperature gradient dT/dr using (17) and (18). When this is done, it is found that $\Delta\nabla T$, the excess over the adiabatic gradient is very small, of the order of one millionth of the adiabatic gradient itself, or of the order of only about 10^{o} K/cm. A typical convective velocity v = 30 cm/sec which is several orders of magnitude smaller than thermal velocities, insuring that convective flow does not upset overall hydrostatic equilibrium. The convective heat flow still contains a parameter, the mixing length, which is difficult to evaluate. Schwarzschild and others have taken ℓ_m to be of order 0.1 R, but this is a problem we need not pursue here. Before leaving this brief discussion of convection, it should be remarked that the convection is turbulent and serves to mix thoroughly the contents of any convective region in a star so the region is pratically uniform in composition.

Since $\Delta\nabla T$ is so small compared with the adiabatic gradient,

$$\left(\frac{dT}{dr}\right)_{ad} = \left(1 - \frac{1}{\gamma}\right) \frac{T}{P} \frac{dP}{dr} \tag{19}$$

the latter can be used in place of $(dT/dr)_{rad}$ from (15) wherever test (16) shows convection to be present.

The four (or five, if we include (19)) structure equations have been written in terms of P, ρ, and T, but these are related by the equation of state of an ideal gas,

$$P = \frac{\rho k T}{\mu H} \tag{20}$$

which is the way astrophysicists customarily write the more familiar form PV = nRT. In (20) k is the usual Boltzmann constant, 1.38×10^{-16} ergs/°K, μ is the mean molecular (or atomic) weight of the gas, and H is the mass of a proton, 1.67×10^{-24} gm. Solving (20) for ρ and substituting into equations (8), (10), (11), and (15) serves to eliminate ρ giving

(i) $$\frac{dP}{dr} = -\frac{GH\mu}{k} \frac{M(r)}{r^2} \frac{P}{T}$$

(ii) $$\frac{dM(r)}{dr} = \frac{4\pi H}{k} \mu r^2 \frac{P}{T}$$

(iii) $$\frac{dL(r)}{dr} = \frac{4\pi H}{k} \mu \varepsilon \frac{P}{T} \tag{21}$$

(iv) $$\left(\frac{dT}{dr}\right)_{rad} = \frac{-3H}{16\pi ack} \mu\kappa \frac{L(r)}{r^2} \frac{P}{T^4}$$

(v) $$\left(\frac{dT}{dr}\right)_{conv} = \left(1 - \frac{1}{\gamma}\right) \frac{T}{P} \frac{dP}{dr}$$

These equations, together with a knowledge of κ and ϵ and their dependence on P, T, μ, and with suitable boundary conditions enable us to construct stellar models.

The boundary conditions will be discussed first. They can generally be cast in a very simple form. While it is true that the stellar surface is not sharply defined, the density does fall to zero rapidly beyond the photosphere in a small fraction of the star's radius R. Similarly the temperature and pressure fall to zero rapidly beyond R, so for our purposes it is permissible to take as boundary conditions:

$$\left.\begin{aligned} P &= \rho = T = 0 \\ M(r) &= M \\ L(r) &= L \end{aligned}\right\} \text{ at } r = R$$

$$\text{and } \left.\begin{aligned} M(r) &= 0 \\ L(r) &= 0 \\ P &= P_c \\ \rho &= \rho_c \\ T &= T_c \end{aligned}\right\} \text{ at } r = 0 \qquad (22)$$

where the subscripts c refer to values at the center of the star. The conditions $M(r) = L(r) = 0$ follow from the definitions of $M(r)$ and $L(r)$. Since $M(r)$ is the mass within a spherical surface of radius r, it clearly vanishes at $r = 0$. Likewise at $r = 0$ there can be nothing left to generate energy so $L(0) = 0$.

The mean atomic weight μ is often more conveniently expressed in terms of the composition. The convection in stellar interior work is to express composition in terms of three variables X, Y, Z which are respectively the mass fractions of H, He, and all heavier atoms in a sample of stellar matter. It follows that $X + Y + Z = 1$ so only two of these variables are needed to specify the composition.

The reason why specifying only H, He, and "heavier" atoms is adequate in most stellar interior work is twofold: (a) in normal, main sequence stars the nuclear energy generating processes represented by ϵ involve conversion of H into He and occasionally conversion of He into C^{12} (i.e., $X \rightarrow Y$ and $Y \rightarrow Z$). The detailed nature of Z is needed only if one wishes to study the processes of nucleosynthesis by which all the heavier elements are built up in stellar nuclear reactions; and (b) as a consequence of the high temperatures in stellar interiors the atoms are fully ionized, for the most part, and the heavier atoms present in appreciable amounts (generally less than 3% by weight) may be considered completely ionized into bare nuclei plus free electrons equal in number to the atomic number of the heavy nucleus, namely $\approx A/2$ electrons for elements C^{12} and up which are the ones of interest.

The molecular weight of a gas is the weight of $N_A = 6 \times 10^{23}$ particles. If 6×10^{23} atoms of H^1 (1 gram) are fully ionized, we have 12×10^{23} particles and 1 gram atom of these (6×10^{23} particles) weighs 0.5 gram since on average half will be protons and half electrons. The mean molecular weight of ionized hydrogen is therefore $\frac{1}{2}$. Similarly a gram-atom of He^4 (4 grams) when fully ionized gives $3 \times 6 \times 10^{23}$ particles (each He^4 gives 1 alpha particle plus 2 electrons); thus the mean molecular weight of fully ionized He^4 is $\frac{4}{3}$. Finally, heavy atoms of normal atomic weight A give rise to A/2 electrons per nucleus, and the mean molecular weight of the fully ionized matter is approximately $\frac{2A}{A+1} \approx 2$. One gram of a mixture consisting of X gms of hydrogen, Y grams of helium, and Z grams of heavy "atoms", fully ionized, produces $2\frac{X}{H}$ particles from H^1 plus $\frac{3Y}{4H}$ particles from He^4 plus $\frac{(A/2)}{AH}Z$ particles from heavier atoms where H is the mass in grams of a single hydrogen atom. The mean molecular weight of this gram of matter is by definition the weight of 6×10^{23} particles.

Since $\frac{1}{H} = 6 \times 10^{23}$, it follows that μ is 1/H divided by the sum of particles given above, or

$$\mu = \frac{1}{2X + \frac{3Y}{4} + \frac{Z}{2}} \qquad (23)$$

In terms of X, Z,

$$\mu = \frac{4}{5X - Z + 3} \qquad (24)$$

The range of μ is from ½ for pure hydrogen up to 2 for heavy matter (Z = 1). As $H^1 \rightarrow He^4$ in the deep interior, X decreases and μ increases, raising each gradient dP/dr, dM(r)/dr, dT/dr, and dL(r)/dr in the basic structure equations. It is not surprising that the increase in μ that results from conversion of H^1 into He^4 has large effects upon the structure of the star.

Consider the next constitutive variable, the opacity, κ. It is clear from the structure equations, (15) or (21,iv), for dT/dr that κ controls the temperature gradient where radiative transfer predominates. In regions of large opacity, the energy cannot escape easily. As a result, the local pressure and temperature rise, causing the star to expand until the energy can get through the high opacity region at just the rate at which it is produced in the interior. It is clear, therefore, that the structure of the star will depend in a sensitive way upon the opacity, for if κ changes, the star has to readjust itself throughout in such a way as to allow the energy generated in the deep interior to get out without being blocked anywhere. The opacity is expected to depend upon P, T, and composition and will vary widely between the center and surface of a star.

Accurate values of κ as functions of the variables are too complex to be represented by a simple analytical formula. For detailed structure work, values of κ are tabulated as functions of P, T, X, Z, and used in that form in computer calculated stellar models. However, analytical representations can be obtained which are sufficiently accurate for approximate structure determinations and which have a further advantage in showing clearly the functional dependence of κ upon T, P, and composition.

Except near the surface, temperatures are so high in normal stars that the atoms are mostly fully ionized, and the stellar material consists mainly of bare nuclei, chiefly protons and helium nuclei, plus enough electrons to give electrical neutrality. These electrons may be free (not bound to any nucleus) or bound temporarily to a nucleus until a sufficiently energetic photon comes along to free it by photoionization. Such bound-free transitions constitute one of the most important processes contributing to opacity. In fact, bound-free transitions ordinarily contribute the dominant term in κ.

Next in importance are free-free transitions. An electron in an unbound state in the vicinity of a positively charged nucleus travels in a hyperbolic orbit. If it encounters a photon, it can be shifted to a different unbound (hyperbolic) orbit with absorption of energy from the photon. Such a process can only take place near another interacting particle (the positively charged nucleus in this case) in order for momentum to be conserved. Finally, free electrons can scatter electromagnetic radiation by the Thomson scattering process. In the cooler stellar envelope, the more familiar bound-bound transitions can take place, giving rise to more or less sharp absorption lines whose wavelengths are characteristic of the absorbing atom. But this process is of negligible importance in the interior.

The bound-free and free-free transitions are much more important than electron scattering in most stars, and it is possible to express the b-f and f-f contributions to κ analytically by a formula first obtained by Kramers. We will not derive the Kramers opacity formula but will simply present it and mention the physical origin of the various factors that appear in it.

$$\kappa_{bf} \approx 4 \times 10^{25} \left(\frac{g}{t}\right) \frac{\rho}{T^{3.5}} \quad (1 + X)\, Z \text{ cm}^{-1}$$

$$\kappa_{ff} \approx 4 \times 10^{22}\, g \frac{\rho}{T^{3.5}} (1-Z)\ (1 + X) \text{ cm}^{-1} \qquad (25)$$

The numerical factors are made up of physical constants such as e, h, m_e, c, etc. The opacity would be expected to be proportional to the density of the gas, ρ, since absorption is certainly proportional to the number of absorbers in the light path. The temperature factor enters in a different way. One may say that at higher temperature the radiation pressure is higher so the radiation flows more easily through the matter leading to reduced opacity. Or one may argue that as T increases, b-f absorption must decrease because the chance is smaller for a given atom to have an electron temporarily in a bound state at any instant. The f-f absorption would also be expected to be smaller at higher temperature where the electron kinetic energies are higher, the hyperbolic paths are straighter, and the electrons are farther from the nuclei on the average, giving a reduced interaction necessary to conserve momentum in the process. All of these ways of looking at the effect of T lead to the expectation of an inverse relation between κ and T. The calculations of Kramers show the dependence to vary approximately as $T^{-3.5}$. The Gaunt factor, g, arises from the wave nature of the electron and is nearly unity. The "guillotine factor", t, was introduced by Eddington to take into account the fact that once an atom has undergone a bound-free transition, it is for a time unable to repeat this process until it later acquires a replacement bound electron again. This effect acts to reduce the absorption below that which would exist if every atom were available for b-f absorption at all times. The guillotine factor depends upon the amount and relative abundances of the heavy elements, upon the density, and upon the temperature in a complicated manner, so the values of t are usually tabulated. Simpler approximate analytical formulas for t are sometimes used in less detailed calculations. For our purposes it is sufficient to take g and t both of order unity.

At low densities and high temperatures scattering of radiation by free electrons (Thompson scattering) can dominate over b-f and f-f transitions which require close proximity between electrons and ions. Thompson scattering gives a simple contribution to the opacity.

$$\kappa_e \simeq 0.19\ (1 + X) \qquad (26)$$

This is proportional to the number of free electrons, $\frac{1}{2}(1 + X)$, as would be expected. The number of electrons is simply the sum of contributions,

$\frac{X\rho}{H} + \frac{2Y\rho}{4H} + \frac{1}{2} A \frac{Z\rho}{AH} = \frac{1}{2}(1 + X)\frac{\rho}{H}$, of electrons from X gms of hydrogen, Y gms of He, and Z gms of heavy elements . Remember that H in the denominator of the summation is the mass of a proton or the reciprocal of Avagadro's number .

Figure 2 below shows the ranges of T and ρ in which the various kinds of opacity dominate.

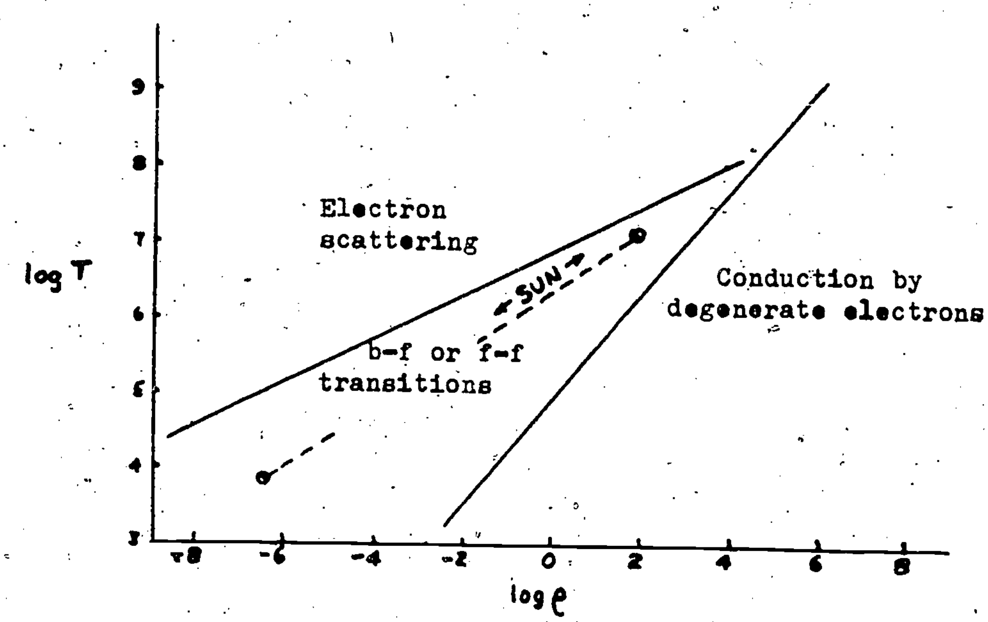


Figure 2. Temperature density diagram for opacity. The run of ρ, T for the Sun is shown.

This figure includes a form of conduction that can be important in regions of very high densities where the electrons are packed so closely that they become degenerate as they are in ordinary metals at low temperatures. Under these conditions heat can be transferred by conduction via the degenerate electrons in much the same way as heat is conducted in a metal. This process is very efficient and leads to nearly isothermal conditions throughout the degenerate region. Electron degeneracy is commonly found in white dwarf star in the cores of low mass stars late in life, and in massive red giant cores after conversion of H^1 to He^4 has ceased in the core.

In ordinary main-sequence stars there are two main processes by which H^1 is fused into He^4. In stars less massive than about 1.2 $M_\odot$ central temperatur and densities favor the proton-proton chain whose principal reaction steps are given below with the energy released in each reaction:

$$\begin{aligned} &H^1 + H^1 \longrightarrow D^2 + e^+ + \nu && 1.44 \text{ Mev} \quad (1.4 \times 10^{10} \text{ yr}) \\ &D^2 + H^1 \longrightarrow He^3 + \nu && 5.49 \text{ Mev} \quad (6 \text{ sec}) \qquad (27) \\ &He^3 + He^3 \longrightarrow He^4 + 2H^1 + \nu && 12.85 \text{ Mev} \quad (10^6 \text{ yr}) \end{aligned}$$

e^+ is a positron and ν is a neutrino. The times in parenthesis are the mean times between reactions for conditions in the center of the Sun. For example, a proton will live on the average for 1.4×10^{10} yrs before it undergoes a collision with another to form a deuteron in the first reaction. The D^2 has a mean life of only 6 sec before it is converted to He^3, and so on. There are other branch reactions in the set, but the three shown above are the main ones. These generate a total of 26.7 Mev of energy per He^4 formed (2 each of the firs two reactions must take place to provide the two He^3 atoms for the final reacti Each neutrino carries away .26 Mev, leaving 26.2 Mev per He^4 to contribute to the star's luminosity.

In more massive stars in which central temperatures exceed $\approx 15 \times 10^6$ °K the following cycle of reactions becomes the dominant energy generating process. This set of reactions comprise the carbon cycle, sometimes known as the CNO bi-cycle since C, N, and O are all involved.

$C^{12} + H^1 \longrightarrow N^{13} + \nu$	1.94 Mev	$(1.3 \times 10^7$ yr)	(28)
$N^{13} \longrightarrow C^{13} + e^+ + \nu$	2.22 Mev	(7 min)	
$C^{13} + H^1 \longrightarrow N^{14} + \nu$	7.55 Mev	$(2.7 \times 10^6$ yr)	
$N^{14} + H^1 \longrightarrow O^{15} + \nu$	7.29 Mev	$(3.2 \times 10^8$ yr)	
$O^{15} \longrightarrow N^{15} + e^+ + \nu$	2.76 Mev	(82 sec)	
$N^{15} + H^1 \longrightarrow C^{12} + He^4 + \nu$	4.97 Mev	$(1.1 \times 10^5$ yr)	

About one in 2500 of the last reactions produces an O^{16} and gives rise to the following set of reactions which are not very important from an energy production standpoint but do illustrate the first steps toward building heavier elements:

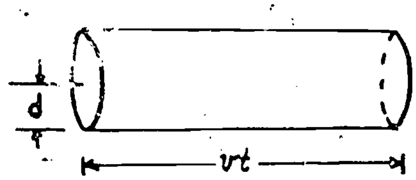
$N^{15} + H^1 \longrightarrow O^{16} + \nu$	(29)
$O^{16} + H^1 \longrightarrow F^{17} + \nu$	
$F^{17} \longrightarrow O^{17} + e^+ + \nu$	
$O^{17} + H^1 \longrightarrow N^{14} + He^4$	

The carbon cycle uses C^{12} as a catalyst since it eventually regenerates the C^{12} except on the rare occasions when the second set of reactions takes place. The main reaction is again fusion of four protons to form a He^4 nucleus. Energy released in this reaction is 25.0 Mev, a little less than in the p-p chain because more energy is lost through neutrino emission in the carbon cycle whose neutrinos are more energetic. The mean reaction times in (28) are again for solar central conditions and vary widely with temperature and density. But they give a rough indication of the slow, overall rate determining steps.

A third kind of nuclear energy generation process will be mentioned briefly the triple-alpha process by which 3 He^4 nuclei fuse to form C^{12} in a two-step process with Be^8 as the intermediate product. This process only goes at very high temperatures, above 10^8 °K, because the strong coulomb repulsion force between two He^4 nuclei requires a large thermal energy to overcome it during a collision in order that the nuclei can approach close enough for the strong, short range, attractive nuclear forces to operate.

The rate of energy generation, ε, by any of the above processes must contain several factors which can be justified by simple physical arguments. First, we note that collisions must occur between two nuclei, such as two protons, if a nuclear reaction is to occur. In fact, these collisions must bring the nuclei within the range of the nuclear forces ($\approx 10^{-13}$cm) for a reaction to occur. It follows that one of the factors in ε must be the number of times per second, on the average, that one kind of nucleus collides with another.

Let the distance between nuclei for interaction via nuclear forces be denoted by d. This may be thought of as the radius of the "sphere of influence" of a given nucleus; if it comes within a distance d of another with which it is capable of reacting, a reaction may take place. As a nucleus of one kind, say type 1, moves about at velocity v, it sweeps out a "cylinder of influence" of volume $\pi d^2 v$ each second.



If there are N_2 nuclei of type 2 per unit volume (with which type 1 can react), our moving type 1 will encounter $\pi d^2 N_2 v$ type 2 nuclei per second, and will pass these closely enough to cause a reaction. If the particle density of type 1 nuclei is N_1 per unit volume, then the total reaction rate must be proportional to $\pi d^2 v N_1 N_2$. This can be converted into mass densities by noting that N_1 is proportional to ρX, and N_2 is proportional to ρX_2 where X_1 and X_2 are the respective abundances by weight. Thus the reaction rate will be proportional to $\pi d^2 \rho^2 v X_1 X_2$. This can be further simplified by taking into account the wave properties of matter. The deBroglie wavelength of a nucleon is h/mv, and the collision distance, d, must be proportional to the deBroglie wavelength. Thus d is proportional to v^{-1} and the collision factor becomes simply $\frac{\rho^2}{v} X_1 X_2$.

So far only nuclei moving with a single relative velocity, v, have been considered. Actually, a distribution of velocities is present, and at a temperature T, the Maxwell-Boltzmann formula for the distribution of velocities in a gas says that the number of nuclei that have relative velocities, v, is proportional to $v^2 T^{-3/2} e^{\frac{-\mu v^2}{2kT}}$ where μ is the reduced mass of the colliding nuclei ($\mu^{-1} = m_1^{-1} + m_2^{-1}$) and k is the Boltzmann constant. This factor must also appear in the final formula for ε.

The colliding nuclei both have positive charges and repel each other electrostatically. In order to get within range of nuclear forces they must penetrate this Coulomb barrier. The Coulomb repulsion is very strong at distance of the order 10^{-13}cm, and it would require very high temperatures to provide enough thermal energy to drive two nuclei this close together. These temperatures would have to be far above those in star centers were it not for the fact that the wave nature of the nuclei permits them to tunnel through a Coulomb barrier that they could not otherwise energetically surmount. Figure 4 illustrates this tunnelling process schematically.

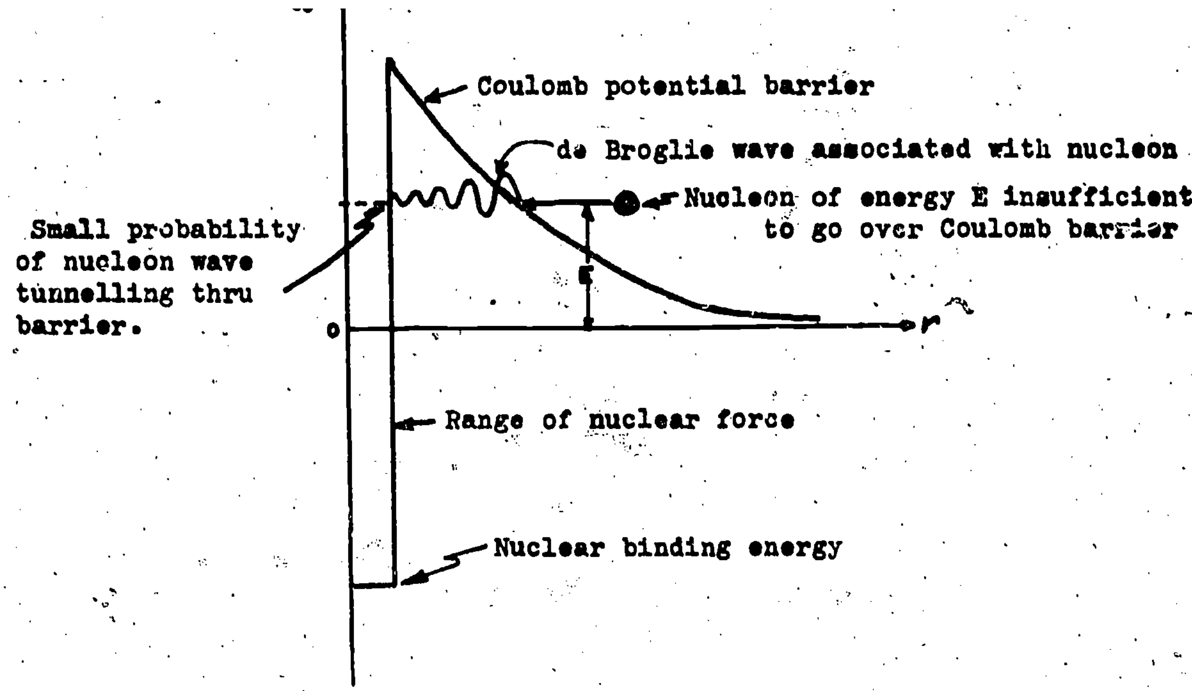


Figure 4. Schematic representation of nucleon tunnelling through Coulomb barrier.

The tunnelling probability, first calculated by Gamow, reduces the number of effective collisions by a factor proportional to

$$\exp\left(-4\pi^2 \frac{Z_1 Z_2}{hv}\right)$$

where Z_1 and Z_2 are the charges on the colliding nuclei, and h is Planck's constant.

Finally, even if two nuclei collide and penetrate the Coulomb barrier, a nuclear reaction will not always take place. To account for this, another factor Γ, the "nuclear factor" must be included. It is the probability that a reaction will take place between two nuclei once they are within the range of nuclear forces. It depends weakly upon the relative velocities of the nuclei and strongly upon the particular nuclear species that are colliding.

When all the factors are put together, the energy production rate, ε, is proportional to

$$\frac{\rho^2}{v} X_1 X_2 \times \frac{v^2}{T^{3/2}} e^{\left(\frac{-\mu v^2}{2kT}\right)} \times e^{\left(\frac{-4\pi^2 Z_1 Z_2}{hv}\right)} \times \Gamma \tag{30}$$

collision factor — Maxwell-Boltzmann factor — Gamow factor — Nuclear factor

This expresssion involving exponential terms that depend upon T can be approximated by much more convenient forms,

$$\varepsilon_{PP} = A\rho X^2 T^n \tag{31}$$

$$\varepsilon_{CNO} = B\rho X X_{CN} T^m \tag{32}$$

where A and B are quasi-constants, (that is, they change slowly with T), m and n are exponents that also change slowly with T, and X_{CN} is the mass fraction of heavier material that consists of carbon and nitrogen. X_{CN} is roughly $\frac{1}{3}$ Z for solar type stars.

In (31) the exponent n is ≈6 at T ≈4 x 10^6 °K, dropping to ≈3.5 at high temperatures around 2 x 10^7 °K. The value of m in (32) also varies from ≈20 at 13 x 10^6 °K down to ≈13 at 50 x 10^6 °K.

In the range of temperatures found in the energy producing core of the Sun, (i.e., T ≈12 to 15 million degrees) n ≈4 and m ≈ 20 while A ≈ 10^{-29} ergs/sec-gm and B ≈ 1.6 x 10^{-142} ergs/sec-gm. Taking $Z_{CN} = \frac{1}{3}$ Z the energy generation equations for a solar type star become, approximately,

$$\varepsilon_{PP} = 10^{-29} \rho X^2 T^4 \text{ ergs/sec-gm} \tag{33}$$

$$\varepsilon_{CNO} = 5 \times 10^{-143} \rho Z X T^{20} \text{ erg/sec-gm} \tag{34}$$

The much stronger temperature dependence of ε_{CNO} reflects the higher Coulomb barrier of the heavy nuclei whose charges are ≈A/2, A being their atomic number, 6 to 8. For completeness we add the approximate expression for $\varepsilon_{3\alpha}$ for the triple alpha process as given by Schwarzschild(5):

$$\varepsilon_{3\alpha} = 10^{-248} \rho^2 Y^3 T^{30} \tag{35}$$

which is valid for $T \approx 140$ million degrees. Y is of course the relative abundance of He^4.

Figure 5 below gives the dependences of the three ε's upon T for the particular choices, $\rho X^2 = 100$, $Z = .015X$, and $\rho^2 Y^3 = 10^8$ for the triple-alpha process:

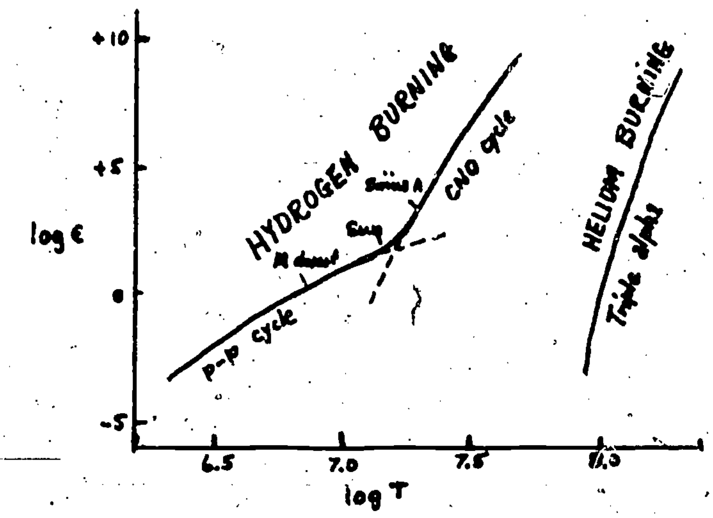


Figure 5. Nuclear energy generation rates as functions of temperature.

The energies released per gram of transmuted matter (i.e., per gram of He^4 or C^{12} formed) for the three processes are

$$
\begin{aligned}
E_{PP} &= 6.3 \times 10^{18} \text{ erg/gm } He^4 \text{ formed} \\
E_{CNO} &= 6.0 \times 10^{18} \text{ erg/gm } He^4 \text{ formed} \qquad (36) \\
E_{3\alpha} &= 6 \times 10^{17} \text{ erg/gm } C^{12} \text{ formed}
\end{aligned}
$$

With these energy release values and the ε's, it is straightforward to calculate the rate at which H^1 is converted into He^4:

$$
\begin{aligned}
\frac{dX}{dt} &= -\frac{\varepsilon_{PP}}{E_{PP}} - \frac{\varepsilon_{CNO}}{E_{CNO}} \qquad (37) \\
\frac{dY}{dt} &= -\frac{dX}{dt} - \frac{3\varepsilon_{3\alpha}}{E_{3\alpha}}
\end{aligned}
$$

These equations, as Schwarzschild points out, contain the core of stellar evolution: the changes in chemical composition produced by nuclear reactions within a star. The structure must change in response to changes in X, Y according to the Vogt-Russell theorem which says that the structure is determined solely by mass (fixed in a given star) and composition which changes with time.

Now that ε has been found in terms of ρ, T, X, Z, it can be expressed with the help of the gas law $\rho = \frac{\mu HP}{kT}$ in terms of P, T. X, Z and inserted into the structure equation (21 iii). All of the structure equations then depend upon just 6 unknowns, P, T, L(r), X, Z. If X and Z are known or inferred from spectroscopic measurements of the outer layers of the star the number of unknowns is reduced to 4 and equals the number of equations (21).

Since the four equations only determine changes in P, T, L(r), M(r) with radius, the procedure for solving the equations involves assuming values of P, T, L(r), M(r) somewhere in the star to start with, usually the center or the surface, and using the equations (21) to calculate values at slightly different r. These values are then used to progressively calculate values at another nearby r, and so on throughout the star.

IV. Calculation of a Simple Stellar Model

To illustrate the method by a simple example, assume μ is constant throughout the star. Such a star would have to be very young, at an age before any appreciable fraction of its hydrogen has been converted into helium. If we take X = 1, Y = Z = 0, we have the simplest possible case.

Next, choose suitable values of P_c, T_c at the center where L(r) and M(r) = 0. With these values we will integrate the structure equations outward, dividing the radius of the star into a number of thin shells of thickness Δr. A finer division would give greater accuracy. There is no reason why all the shells should be of the same thickness, but division into equal Δr will be used in this case for illustrative purposes. We arbitrarily pick starting values of

$T_c = 8 \times 10^6\ {}^oK$ and $P_c = 8 \times 10^{16}$ dynes/cm^2 because these lie in the general range of values for a star somewhat less than one solar mass. The gas law gives

$$\rho_c = \frac{\mu H P_c}{k T_c} = \frac{.5 \times 1.67 \times 10^{-24} \times 8 \times 10^{16}}{1.38 \times 10^{-16} \times 8 \times 10^6} = 60.5 \text{ gm/cm}^3.$$

Any of the four variables P, T, L(r), $\mathcal{M}(r)$ at r can be calculated at $r + \Delta r$ by means of a Taylor's series expansion about r. For example,

$$P(r + \Delta r) = P(r) + \frac{dP}{dr}\Big|_r \Delta r + \frac{1}{2}\frac{d^2P}{dr^2}\Big|_r \Delta r + \ldots\ldots$$

We will take only the first derivatives (from equations 21) except where they vanish at r=0. In the present example Δr will be chosen to be 10^9cm, small enough to give a fairly smooth run of the variables P, T, etc.

In the first region, at $r = 10^9$cm, designated r_1,

$$M(r_1) = \frac{4\pi}{3} r_1^3 \rho_c = 2 \times 10^{30} \text{ gm}.$$

Since at the center $\left(\frac{dP}{dr}\right)_c = 0$, it is necessary to take the second order term in the series expansion:

$$\left(\frac{d^2P}{dr^2}\right)_c = -\frac{4}{3}\pi G\left[2\rho\frac{d\rho}{dr}r + \rho^2\right]_c = -\frac{4}{3}\pi G\rho_c^2 \text{ because } \left(\frac{d\rho}{dr}\right)_c = 0.$$

Taking $G = 6.67 \times 10^{-8}$, its value in cgs units, and $\rho_c = 60.5$, we find

$$P(r_1) = P_c - \frac{1}{2}\left(\frac{4}{3}\pi G\rho_c^2\right) r_1^2 = 8 \times 10^{16} - 2 \times 10^{15} = 7.8 \times 10^{16} \text{ dynes/cm}^2.$$

The luminosity at r_1 becomes

$$L(r_1) = \left(\frac{dL(r)}{dr}\right)_c \times r_1 + \frac{1}{2}\left(\frac{d^2L}{dr^2}\right)_c r_1^2$$

By (11) $(dL/dr)_c = 0$ since it is proportional to r^2. Again the 2nd derivative must be used:

$$\left.\frac{d^2L}{dr^2}\right|_c = 4\pi \left[2r\rho\varepsilon + r^2\varepsilon \frac{d\rho}{dr} + r^2\rho \frac{d\varepsilon}{dr}\right]_c = 0$$

Therefore a third term in the expansion is needed. It is

$$L(r_1) = \frac{1}{3!}\left(\frac{d^3L}{dr^3}\right)_c (\Delta r)^3 = \frac{2\pi}{3}\left[2\rho\varepsilon + \text{terms proportional to } r, r^2, \text{etc.}\right] \Delta r$$

$$= \frac{4\pi}{3}\rho_c \, \varepsilon_c \, r_1^3$$

where we have used $\Delta r = r_1 = 10^9$cm, as before.

This star has a relatively low T_c and the p-p chain will be the only important mode of energy production. Therefore, we can take for ε at the center from Eqn. (33),

$$\varepsilon_c = (\varepsilon_{pp})_c = 10^{-29}\,\rho_c\,T_c^{\,4}$$

Taking $\rho_c = 60.5$ gm/cm^3 and $T_c = 8 \times 10^6$ ^oK, we obtain $\varepsilon_c = 2.48$ ergs/sec-gm, and $L(r_1) = 4.6 \times 10^{30}$ erg/sec.

Since T has a maximum at the center, $(dT/dr)_c = 0$, and a second term in the expansion is required, it is

$$\frac{1}{2}\left(\frac{d^2T}{dr^2}\right)_c (\Delta r)^2 = -\frac{3\,\kappa_c\rho_c}{32\pi acT_c^{\,3}}\left[\frac{1}{r^2}\frac{dL(r)}{dr} - 2\frac{L(r)}{r^3}\right]_c r_1^{\,2} = \frac{-\kappa_c\rho_c^{\,2}\varepsilon_c r_1^{\,2}}{8ac\;T_c^{\,3}}$$

Therefore $T(r_1) = T_c - \dfrac{\kappa_c\rho_c^{\,2}\varepsilon_c r_1^{\,2}}{8ac\;T_c^{\,3}}$

The opacity κ_c must be evaluated before $T(r_1)$ can be calculated. Taking the Kramers value for free-free transitions (b-f transitions may be neglected when Z = 0) and adding the electron scattering opacity (equation 26) gives

$$\kappa = 0.39 + 8 \times 10^{22}\,\frac{\rho_c}{T_c^{3.5}} = 0.39 + 3.34 = 3.73 \text{ cm}^{-1}$$

With this opacity $T(r_1) = 7.86 \times 10^6$ ^oK.

In Table 1 all of the above values at $r = 0$, $r_1 = 2 \times 10^9$cm and at successive $\Delta r = 2 \times 10^9$cm intervals are listed. The applicable structure formulas are collected below in numerical form

$$\Delta M(r) = 12.6\rho r^2 \Delta r$$

$$\Delta P(r) = -6.67 \times 10^{-8} \rho \frac{M(r)}{r^2} \Delta r$$

$$\Delta T = -263 \frac{\kappa\rho L(r)}{r^2 T^3} \Delta r \text{ (radiative transport)}$$

$$\Delta T = -1.62 \times 10^{-16} \frac{M(r)}{r^2} \Delta r \text{ (convective transport)}$$

$$\Delta L = 12.6\, r^2 \rho \varepsilon \Delta r \qquad (38)$$

$$\varepsilon = 10^{-29} \rho T^4$$

$$\kappa = .39 + 8 \times 10^{22} \frac{\rho}{T^{3.5}}$$

$$\rho = 6.05 \times 10^{-9} \frac{P}{T}$$

$$\left|\frac{\Delta T}{T\Delta r}\right| < \left|0.4 \frac{\Delta P}{P\Delta r}\right| \text{ condition for stability against convection}$$

$$\rho = 8.95 \times 10^{-9} P^{3/5} = 1.30 \times 10^{18} T^{-2.5} \text{ (convective transpo}$$

(NOTE: The formula for ΔT (convective) comes from $\frac{-\Delta T}{\Delta r} = (1 - \frac{1}{\gamma}) \frac{\Delta P}{\Delta r}$ with $\gamma = 1.66$ which is appropriate for a monatomic gas. Replacing $\frac{\Delta P}{\Delta r}$ by $-\frac{GM(r)\rho}{r^2}$ and noting that $\rho = \frac{\mu HP}{kT}$ gives the formula shown. The relations between ρ and P for convective (adiabatic) transport is the familiar adiabatic gas law $\rho = KP^{\frac{1}{\gamma}} = K'T^{\left(\frac{1}{\gamma-1}\right)}$ with K, K' constants which are evaluated from particular values of (ρ, P) and (ρ, T) calculated in the convective region of Table 1 below.)

TABLE 1: A pure hydrogen model star integrated from r = 0 outward.

r (10^{10}cm)	M(r) (10^{33} gm)	P(r) (10^{15} dynes/cm^2)	T(r) (10^6 °K)	L(r) (10^{33} erg/sec)	ρ (gm/cm^3)	ε (erg/sec-gm)	κ (cm^{-1})
0	0	80	8.0	0	60.5	2.48	3.73
.2	.002	78	7.86	.0046	60.0	2.29	3.92
.4	.008	74	7.57	.014	59.1	1.93	4.35
.6	.0104	70	7.30	.060	58.0	1.65	4.81
.8	.063	67.8	7.00	.147	56.9	1.54	4.94
1.0	.109	60.3	6.30	.288	57.9	.91	7.78
convection 1.2 begins →	.255	51.9	3.51	.420	88.0	.14	82.3
1.4	.415	41.5	3.00	.465	83.7	.075	143.6
1.6	.828	17.9	2.31	.468	46.9	.013	201
1.8	1.13	-2.3	1.26	.472			

This calculation has been carried only as far as $r = 1.8 \times 10^{10}$cm where P(r) becomes negative. It must go to zero just a bit short of this radius where T should also become zero as required by the surface boundary conditions. The fact that both P and T do not simultaneously become zero for some r which would then be the star's radius, R, is a sign that the correct initial conditions were not chosen at r = 0. This is no surprise since P_c and T_c were chosen arbitrarily for the first run through the calculation. A trial and error approach could be adopted, going back to r = 0 and changing the central parameters P_c, T_c slightly and repeating until T(R) = P(R) = 0 simultaneously. There are more sophisticated methods for insuring a fit at the surface[6], but these will not be described here.

Our model is unsuccessful in the sense that our first guess at initial conditions did not produce a possible star. It did serve to illustrate the numerical integration of the structure equations considered as finite difference equations (38) rather than as differential equations (21). The model also illustrates the development of a convective envelope on a radiative core. Initially at about r = 1.2 to 1.4 a suspiciously large $\Delta T/\Delta r$ was found using the radiative transport equation for $\Delta T/\Delta r$. When this was tested by the stability criterion (Eq. 16), it was found that convection takes place in this region. So $\Delta T/\Delta r$ was recalculated using the convective form of $\Delta T/\Delta r$ from Equation (38). At every succeeding step this test had to be repeated to be sure convection persisted.

Stellar models are generally presented in tabular form as in Table 1 above or in graphical form in which various parameters P(r), T(r), L(r), M(r) are plotted against r. In Figures 6 and 7 below, this is done for our first trial model.

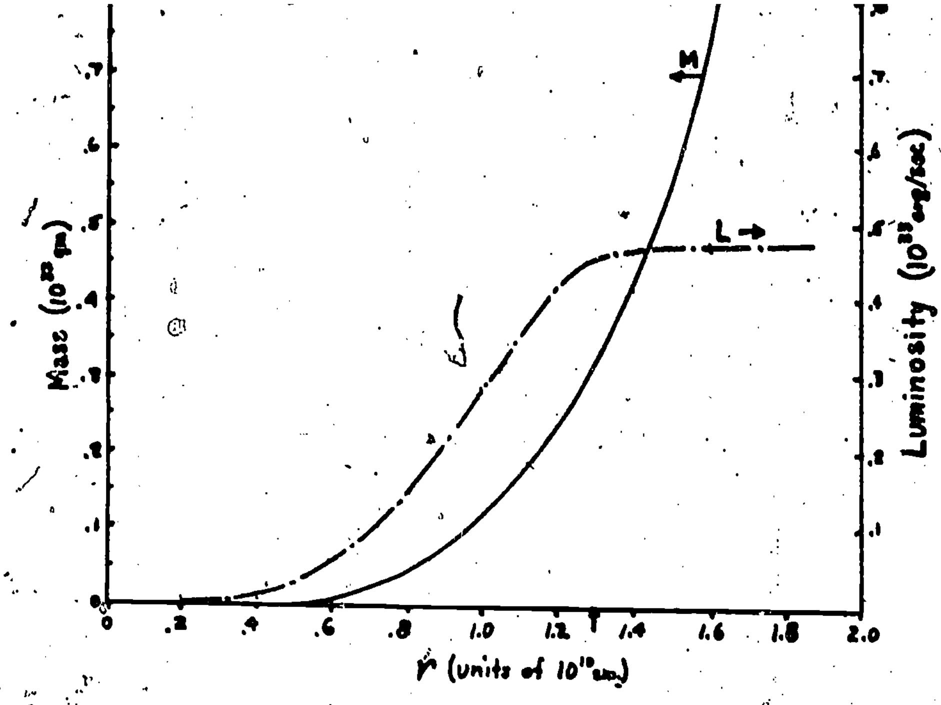


Figure 6: Mass and luminosity variation with radius for trial model star. The arrow on the abscissa shows the position of the boundary between the radiative core and the convective envelope.

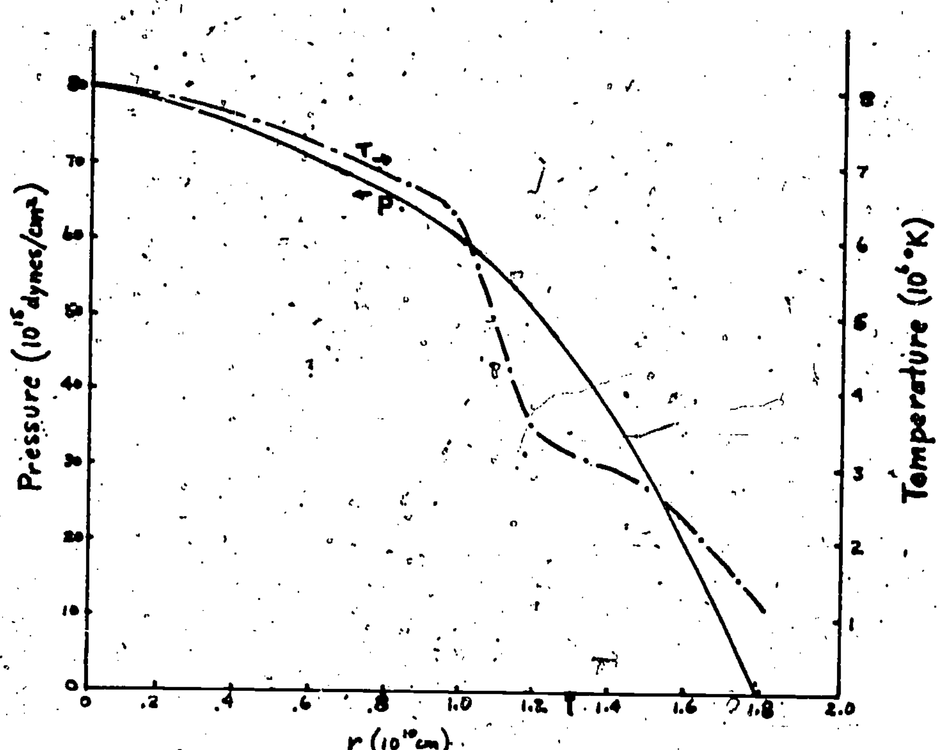


Figure 7: Radial dependence of P, T in a trial model star with radiative core and convective envelope. The arrow at r = 1.3 x 10^{10}cm marks the transition between these regions.

In Figure 7 the boundary between radiative and convective regions is clearly indicated by the marked levelling off of the T(r) function. This is to be expected because convection is an efficient means of energy transport and consequently convective regions have smaller temperature gradients than adjacent radiative regions. In Figure 6 the convective region is seen to lie at about the point where the luminosity L(r) ceases to increase because nuclear energy generation has ceased ($\varepsilon \approx 0$) where the temperature has become too low to sustain appreciable hydrogen burning.

The mass of the trial model cannot be determined until the model has been adjusted to give proper surface conditions, P(R) = T(R) = 0. Figure 6 shows that M(r) is still rising at $r = 1.8 \times 10^{10}$. At this point $M = 1.13 \times 10^{33} \approx 0.57\ M_{\odot}$ so it appears that this model, properly readjusted, would yield a star of $\approx$ 0.6 - 0.65 solar masses. Its luminosity would be $\approx 0.48 \times 10^{33}$ erg/sec or about 1/8 $L_{\odot}$, and its radius would be roughly 2×10^{10} cm, or about 0.3 $R_{\odot}$. With these rough estimates of L and R, the effective surface temperature becomes $\approx$ 80,000°K from the Stefan-Boltzmann equation.

This would be a very blue, hot, sub-dwarf star, well below the main sequence and far to the left, in the region generally associated with the central stars of planetary nebulae.

The general features of the curves in Figures 6 and 7 are repeated in the following figures that illustrate real, detailed model calculations for main sequence stars of various masses.

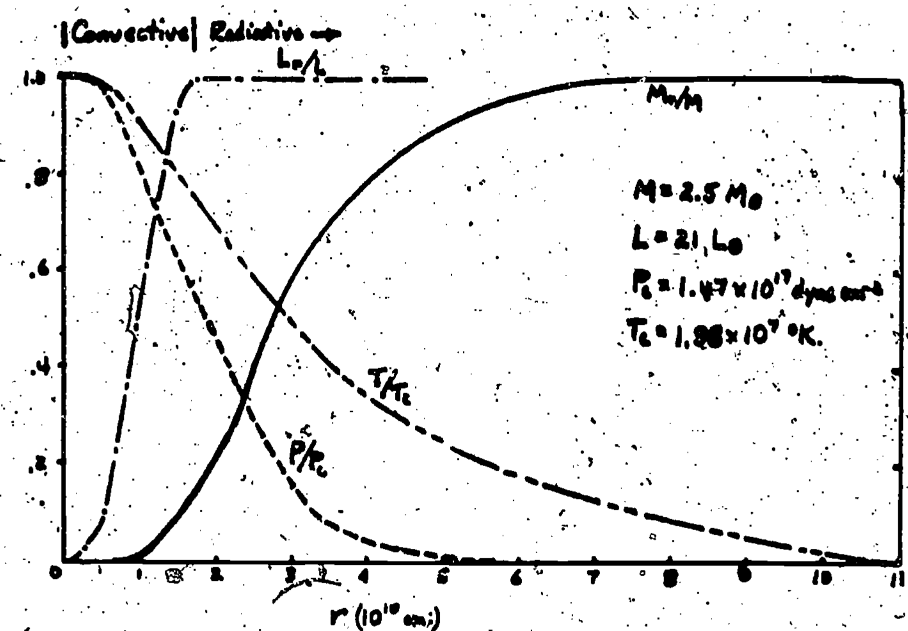


Figure 8: Graphical representation of a model star of 2.5 $M_\odot$, R = 1.59 $R_\odot$ of composition X = 0.90, Y = 0.09, Z = 0.01[6], L = 21 $L_\odot$ for this upper main sequence star. It has a convective core and radiative envelope and energy is generated mainly by the carbon cycle. The boundary between core and envelope is at r ≈ 2 x 10^{10}c

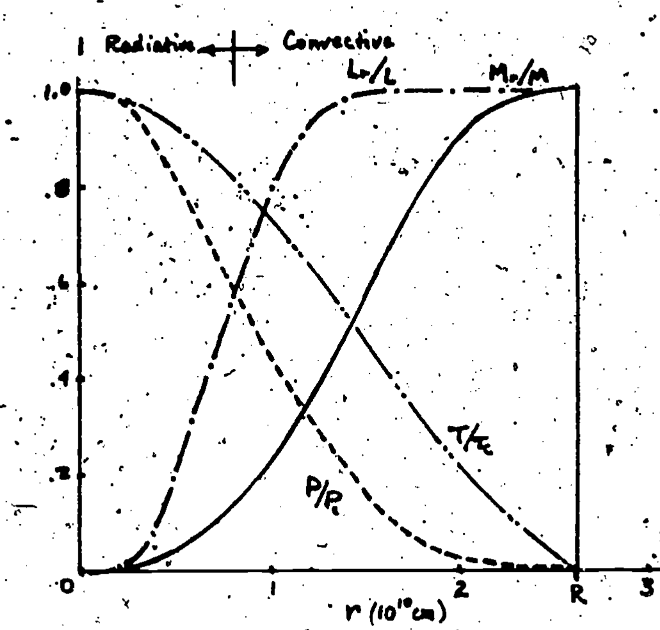


Figure 9: Model star of 0.6 $M_\odot$, R = 0.375 $R_\odot$ of composition X = 0.72, Y = 0.26, Z = 0.02[7], L = 0.4 $L_\odot$ for this star which has a radiative core and a convective envelope joined at r ≈ 0.85 x 10^{10}cm. The star generates its energy by the p-p process. Its surface temperature is 7500°K, somewhat hotter and bluer than the Sun.

The model star shown in Figure 9 is rather like our earlier trial model in size and mass. It is a good, equilibrium star while the trial model was not. The higher central pressure and temperature suggest that the trial model ought to be modified in this direction.

The astrophysics literature contains many examples of stellar models calculated for a wide range of masses and compositions. These, plotted on the H-R diagram and mass-luminosity diagram reproduce quite satisfyingly the distributions of real stars on these diagrams.

The field of stellar evolution which is a generalization of the field of stellar structure determination actually involves the calculation of whole series of models allowing a small composition change between successive models of the kind that would take place in a star converting its internal stores of H^1 into He^4 and heavier elements. These successive models involve a calculable time lapse between them during which the composition change occurs. When the series is connected by a line on the H-R diagram, such a line represents an evolutionary track on the diagram along which a star moves as it ages. The various forms such tracks take depend upon both initial mass and composition of the star, and they form the basis for a discussion of stellar evolution.

REFERENCES

1. J. P. Cox and R. T. Giuli, "Principles of Stellar Structure" Vol. 2, Gordon and Breach, N. Y. (1968), Ch. 18, pp 569-579.

2. E. N. Parker, "Annual Review of Astronomy and Astrophysics" Vol. 8, Annual Reviews, Inc., Palo Alto, Calif., 1970, pp 1-30.

3. J. N. Bahcall and R. Davis, Jr., Science, 191, 264 (1976).
   R. W. Noyes in "Frontiers of Astrophysics" Ed. by E. H. Avrett, Harvard University Press, Cambridge, Mass., (1976), pp 43-46.

4. M. Schwarzschild, "Structure and Evolution of the Stars", Dover, N. Y., Chapter V.

5. ibid, Ch. II, Sec. 10.

6. ibid, Ch. III. See also E. Novotny "Introduction to Stellar Atmospheres and Interiors", Oxford University Press, N. Y., 1973, Part 4.

7. See Novotny, pp 467-486.

PULSARS

I. History

Pulsars were first discovered in late 1967, more or less by accident. A group of radio astronomers at Cambridge, England, (Hewish, Bell, Pilkington, Scott, and Collins)(1) had just put into operation a new radio telescope optimized to detect scintillations in radio sources caused by small scale variations in the interplanatary gas or "solar wind". This was part of an important program to measure the sizes of astronomical radio sources and from size measurements to determine their distances. An important feature of the radio detection system of the telescope was its capability to respond to rapid fluctuations in signal strength on a millisecond time scale. This response had been designed into the system to permit it to record scintillations, but it also gave it the capability to respond to pulsar radio emission.

Shortly after this new telescope went into operation, the Cambridge group began to notice occasional sporadic interference, the strongest appearing to come from a location in the constellation Vulpecula. This source was periodic and consisted of short pulses of 81.5 MH_z energy (the frequency to which the telescope was tuned) repeated precisely every 1.337 seconds. This first pulsar was reported the following February, 1968, in Nature and was confirmed promptly by other groups at Jodrell Bank, Arecibo, CSIRO in Australia, and at JPL in California. Meanwhile Ryle and Bailey(2) made a more precise determination of the location of this pulsing radio source using the Cambridge 1-mile interferometer. They found the radio source to be located near enough to a faint, somewhat reddish star to suggest strongly that the star and radio source were one and the same.

Pilkington and his coworkers(3) rechecked other locations where interference had been found among the scintillation measurements and quickly found three more pulsars. Meanwhile, workers at Kitt Peak began searching for optical pulsations, and early in May Lynds, Maran, and Trumbo(4) found what seemed to be a regular but very small variation in light from the star believed to be associated with the first pulsar, but these light variations had a period twice that of the radio pulses and a different shape.

The search for more of these mysterious objects was on in earnest, and to date more than 100 have been found and studied. Probably the most valuable of these from the standpoint of the help it has provided in understanding the nature of pulsars is the Crab pulsar, NP0532. This object was first discovered in the Crab Nebula in mid-1968, and it has the shortest pulse period, 33 milliseconds, of any pulsar yet found. The following January a group of astronomers(5) in Arizona (Cocke, Disney, and Taylor) made a momentous discovery when they found regular pulsations in the light from a central star in the Crab Nebula long believed to be the remnant of the supernova that produced the nebula in AD 1054. This optical pulsar had exactly the same period as the radio pulsar (.033095 seconds) and was within the region from which the radio signals came. Identification of the Crab pulsar with the Crab supernova remnant was important in two respects: it connected pulsars and supernova residues positively, for the first time, and it allowed theorists to draw upon a large body of information that had accumulated about the Crab Nebula which has been carefully studied because it is the closest recent supernova and the one that can be most fully resolved by our telescopes.

II. Observed Properties of Pulsars

The highly unusual character of the pulsar implied by its rapid and incredibly precise pulse period generated enormous interest immediately following the discovery of the first of these objects. As a result, most of the world's radio telescopes were quickly put to searching for more pulsars and studying those already found. The results of this observational activity have led to the following picture of pulsar characteristics.

The bursts of radio energy that constitute the observed pulses contain radiofrequency energy covering a very wide frequency band, in some cases into the optical, ultraviolet and even x-ray regions of the spectrum. Each pulsar has its individual pulse shape but only when averaged over a large number of pulses. There is considerable variation in shape from pulse to pulse from a given pulsar. Figure 1 shows characteristic pulse shapes from four pulsars. These shapes are averaged over a large number of individual pulses in each case. The shapes also depend somewhat upon the radiofrequency at which the observations are made.

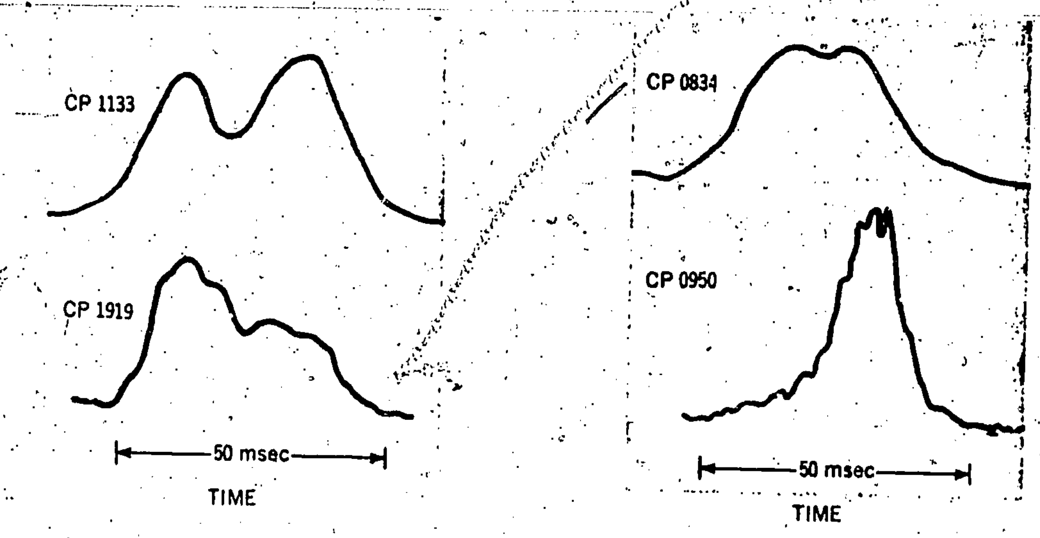


Figure 1. Average pulse shapes of four pulsars. The ordinate is energy flux received on Earth but normalized to the same peak value for each pulsar.

The most striking and significant feature of pulsar emission is the regularity of the pulse repetition rate. These periods range from 33 milliseconds up to 4.8 seconds and the pulse period averaged over a large number of pulses is precise to a few parts per billion, making them among the most precise "clocks" known. Even a laboratory quality quartz crystal oscillator can scarcely match this precision which is in the range of atomic frequency standards.

A second important observation on the pulse rate is that all pulsars are slowing down at measureable rates as determined by a lengthening of the pulse period. The rate of slowing down can be used to deduce a "lifetime" for the pulsar. The observed rate of change of the pulse period has been found proportional to the period, P, a fact which can be expressed

$$\frac{dP}{dt} = -\frac{P}{\tau} \qquad (1)$$

This leads to a time dependence of the form $P(\tau) = P(0)\ e^{-t/\tau}$ which identifies τ as a lifetime, the time it takes for the pulse rate to decrease by e^{-1}. In this way it has been found that pulsar lifetimes are of the order of 10^4 to 10^7 years, rather short times by cosmic standards.

Two brief exceptions to the steady running down of the pulsar clock have been observed. Several pulsars have abruptly exhibited a small increase in pulse rate, and following these "glitches" they have again settled down to resume their normal slowing down rates. We shall later see that this phenomenon has revealed an intriguing feature of pulsar structure.

Detailed analysis of individual pulse structures shows that several things are apparently going on simultaneously in a pulsar, for there are three distinct time scales involved. There is first the long time scale of the pulse rate, namely, .033 to 4.8 seconds. Then there is a shorter time scale, on the order of milliseconds, that accounts for much of the variability in shape from one pulse to another. Finally, the shortest observed time scale is on the order of 100 microseconds, and this time scale sets an upper limit to the size of the region that radiates the pulsar pulses.

The argument goes like this: if a radiator emits periodic electromagnetic waves at a more or less definite frequency, all parts of the radiator that contribute to this radiation must act in phase. If this were not the case, different parts of the body would radiate out of phase with each other and would on the average cancel each other's radiation. Now if all parts of a radiator cooperate to radiate in phase, each part must somehow receive information about the others, otherwise they could never remain synchronized in phase. Relativity theory teaches us that no information can be transmitted faster than the speed of light, c. Therefore, parts of a radiator that are farther apart than an electromagnetic wave can travel in about a period of the radiation frequency cannot stay synchronized because they are effectively out of touch with each other on the time scale of the radiation emitted. This means that a radiator emitting regular pulses at intervals of 100 μ sec must not be larger than light can travel in 100 μ sec. This sets an upper limit on the size of the radiator in a pulsar, and this size limit is about 30 km (10^{-4} sec x 3 x 10^{5} km/sec). This does not necessarily mean that the entire object that constitutes the pulsar cannot be larger than this; the radiator might be a "spot" of this size on the surface of a larger body, and additional information is needed before a choice can be made between these possibilities.

Another important observation is that the radiofrequency energy in the pulses is polarized. Polarized radio waves have the same general properties as polarized light, and radio polarization is recognized by using receiving antennas that respond to polarized incident waves. Polarization in pulsar radiation was found to be very complex. It depends upon frequency and is highly variable, changing from pulse to pulse and even within a single pulse. But the significance of the finding that the radiation is strongly polarized is that it very likely originates in a region in which there is a strong magnetic field because we know that radiation from regions permeated by magnetic fields is characterized by strong polarization. This conclusion is supported by the observed spectral distribution of the radiation and by the fact that the Crab Nebula contains strong fields, as determined by independent observations of other features of the Nebula. Thus it seems firmly established that strong magnetic fields are associated with pulsars.

In order to deduce the amount of energy radiated by a pulsar, its distance must be known. One of the important consequences of identification of the Crab Nebula with NP0532 is that the distance of the Crab is accurately known, and this permits a good determination of the energy radiated by the Crab pulsar.

By one method or another all the known pulsars have been found to lie within our Galaxy. Pulsars undoubtedly exist in other galaxies, but they are too weak to be observable. The distribution of pulsars in the Galaxy is consistent with their identification as supernova remnants. Their number also supports this view. Since they are relatively short lived objects ($< 10^7$ years), there should not be many around to observe at any given time; if they are the consequences of supernova explosions, their number should be correlated with the observed frequency of supernova occurrence, and this turns out to be the case, allowing for a number of observational uncertainties.

The energy loss rate from a young pulsar like the Crab has been found to be of the order of 10^{38} ergs per second. This includes radiation over the entire electromagnetic spectrum and is approximately 10^5 times greater than the radiation from the Sun. If an object radiates at this rate for a lifetime of $\approx 10^7$ years, it must lose in that time a total of $\approx 10^{52}$ ergs. We shall use this figure later to see what pulsar models can contain this much energy.

III. Basic Pulsar Models

We know only three general types of massive objects in nature that are good timekeepers and therefore possible candidates for models of pulsars whose chief characteristic is its remarkable regularity. These classes of objects that are capable of relatively slow periodic motions are vibrating stars, binary stars, and rotating stars. Let us examine each briefly in light of observed pulsar characteristics.

Certain types of stars are known to pulsate as a result of radial oscillations of their atmospheres. The periods range from days to months and the periods do not exhibit the order of precision of pulsar pulses. The general theory of vibrating bodies leads to a dependence of the period P upon the density ρ of the form $P \propto \rho^{\frac{1}{2}}$ or upon mass of the form $P \propto M^{\frac{1}{2}}$. These two simple expressions permit us to rule out vibrating stars as pulsar models because the second predicts the wrong variation of pulse period with time. As time goes on the mass of a star must ordinarily decrease. It loses some mass by matter ejection in the form of a stellar wind, and it radiates away mass in the form of energy ($E = Mc^2$). Consequently if M decreases, the vibration period should decrease and the pulse rate ought to increase rather than decrease, as observed. This argument is not by itself conclusive, however, because stars may also accrete matter under certain circumstances by gravitational attraction of the gaseous envelope from a close binary red giant companion, for example. The accretion rate could exceed the mass loss by radiation and stellar wind and could lead to a pulse rate slowing as observed. But when one calculates the densities required to give pulse rates of a second or less, a problem crops up. The most dense white dwarf stars are not able to vibrate rapidly enough. In fact, unless one postulates rather unlikely complex modes of vibration, likely to be soon damped out, it does not appear possible to devise a plausible vibrating white dwarf model with a period shorter than ≈ 5 seconds. This is close enough to the 4.8 sec. period of the slowest pulsar but is too far from a 33 millisecond period (Crab) to be explained away by uncertainties in the calculation. Only a star with density in the range of nuclear densities (> 10^{11} gm/cm^3) could vibrate as rapidly as the Crab pulsar.

A second general group of candidates for pulsar models is the class of closely spaced binary stars revolving about their common center of mass much like a rotating dumbbell. Let us see what kind of binary would be required to rotate 30 times per second. The laws of orbital motion (Kepler and Newton) tell us that the masses, periods, and sizes of the orbit of the members of a binary pair are related by Kepler's 3rd law

$$(M_1 + M_2) P^2 = 4\pi \; Ga^3 \qquad (2)$$

where G is the gravitational constant and a is the semi major axis of the elliptical trajectory. If the members of a binary are comparable in mass to the Sun, then using P = .033 seconds we find $a \approx 2 \times 10^7$ cm, about $\frac{1}{300}$ the radius of the Earth! Thus a binary model would require two approximately solar mass objects to revolve in orbits much smaller than the radius of the Earth. If these condensed stars were in contact their individual radii could not exceed 2×10^7 cm and their densities would be of the order of $\rho \approx M_{\odot}/\frac{4}{3}\pi a^3$ or about 6×10^{10} gm/cm³. Again we find that a stellar object with a density of the order of 10^{11} gm cm^{-3} or higher is required. This is the lower limit for the density of a neutron star, a body in which nearly all of the matter normally in the form of protons and electrons has been converted into neutrons by the enormous pressures and temperatures involved. The neutrons are formed by the reaction $p^+ + e^- \rightarrow n$, and the protons and electrons come from the matter of the star, mainly hydrogen and helium, that has been completely ionized.

We see that whatever the defects of the vibrational or binary star models may be, both call for objects of densities in the neutron star range. We shall soon see that the rotating object model also demands a neutron star, so the decision is unanimous. But before the rotating model is considered, we must point out two fatal defects of the binary model of a pulsar. The first of these is very simple. If the orbital kinetic energy of the binary is what feeds the pulse energy, as the binary loses energy it must speed up, not slow down as observed. This follows from Kepler's law above and the fact that the orbital size, a, must decrease as energy is extracted from the system. In this respect the binary model has the same drawback as the vibrational model: both should speed up rather than slow down as time goes on. There is another more esoteric defect of the binary model. General relativity theory says that two very massive bodies orbiting close together, as the model demands, are strong radiators of gravitational waves. These waves carry energy away from the binary system, and calculation reveals that a 33 millisecond binary would be such a strong gravitational wave source that it would lose all its energy in a relatively short time, far less than 10^7 years.

Finally, we turn to the third possibility for a pulsar model, a rotating body with a radiating source somehow attached to its surface so the radiation sweeps around in space like a lighthouse beam. The pulses would be received on Earth each time the beam sweeps by, and the pulse rate would give the rotational period of the object.

A simple argument can be used to set an upper limit on the size of such a pulsar model. We require that its surface gravity at least equal the centrifugal force at its equator; otherwise it would lose matter rapidly from the equator. If the total mass and radius of the object are M and R respectively, the gravitational acceleration at the surface is simply

$$g = \frac{GM}{R^2} \qquad (3)$$

where G is the gravitational constant (6.67×10^{-8} cgs units). At the equator the centripetal acceleration is $R\omega^2$ where ω is the angular frequency of rotation. For zero mass loss we require

$$R\omega^2 < \frac{GM}{R^2} \qquad (4)$$

The mass may be written in terms of the average density as $M = \frac{4}{3} \pi R^3 \bar{\rho}$. Also writing $\omega = 2\pi/P$ allows (4) to be put into the more simple form

$$\bar{\rho} \gtrsim \frac{3\pi}{GP^2} \simeq \frac{10^8}{P^2} \quad \text{gms/cm}^3 \qquad (5)$$

This establishes the minimum average density for the slowest pulsar at $\simeq 10^7$ gm cm^{-3}, in the white dwarf range, and for the Crab pulsar $\bar{\rho} \simeq 10^{11}$ gmcm^{-3}, well above the WD range and at the lower end of the neutron star range.

Next, the rotational energy can be estimated from elementary physics. The rotational kinetic energy of a body of moment of inertia I is simply

$$E = \frac{1}{2} I\omega^2 \tag{6}$$

and the power dissipated when ω changes is given by

$$\frac{dE}{dt} = I\omega \frac{d\omega}{dt} \tag{7}$$

For the moment of inertia $I \simeq MR^2$ we may take with enough accuracy for an order of magnitude calculation, $M \simeq M_{\odot}$ and $R \simeq 15$ km from considerations above involving light travel time between pulses on a 100 μsec time scale. This gives a total rotational energy $E \simeq 2 \times 10^{47} P^{-2}$ ergs. If P = .033 sec, $E \simeq 2 \times 10^{50}$ ergs. The measured slowing down rate of the Crab pulsar is 4×10^{-13} periods/period or 1.3×10^{-5} seconds per year. At this rate its lifetime should be 4×10^4 years, and if its rotation supplies all the energy of the nebula ($\simeq 3 \times 10^{38}$ ergs/sec), the stored energy ought to be approximately 4×10^{50} ergs in this particular pulsar. This agrees well enough with the estimated rotational energy of the object to suggest strongly that a dense, compact, rotating body is indeed a suitable model for a pulsar. Such a model also slows down with time as it loses rotational energy in agreement with observation. It does not radiate gravitational waves, so this energy loss mechanism does not operate to reduce the lifetime as in the binary star case.

If a rotating neutron star is to be acceptable as a pulsar model, we must be able to account for its rapid rotation which must be equal to the pulse rate or differing from it by a small integer factor (i.e., maybe there are two or four pulses per revolution instead of just one). If the neutron star conserves angular momentum during its formative collapse, its rotation rate can be accounted for. Consider a star like the Sun with $R = 7 \times 10^{10}$ cm and rotational period ≃ 21 days (≃ 2×10^6 sec). If such a body were to contract to $R \simeq 10$ km ≃ 10^6 cm conserving angular momentum ($\propto MR^2\omega$), its rotational period would decrease by a factor $(R_1/R_2)^2$, i.e., by $(7 \times 10^4)^2 = 5 \times 10^9$ in this example. The final period would then be ≃ 0.5 millisecond, more than short enough for a new pulsar. If one takes into account the uncertainties in this crude calculation, the agreement with observation is satisfactory.

It was mentioned earlier that in a few pulsars there have been abrupt changes in period, leading to the conclusion that the pulsar speeds up suddenly and then soon settles down to its former rate of slowing down. These speed-ups or glitches are quite small, but easily measurable. What can cause a rotating object to speed up if no external torque is applied? The answer is simple; its radius decreases causing a decrease in its moment of inertia I; as a consequence, the rotational frequency ω goes up and the period decreases. The mystery about why the neutron star should suddenly decrease its radius by about 1 cm, the amount needed to produce the observed speed-up, was solved by Ruderman(6) who showed that the surface layer of a neutron star was probably solid and in fact, crystalline. Stresses in this crust from centrifugal and Coriolis forces occasionally cause it to crack or to recrystallize, producing a "starquake". The resulting rearrangement of the stellar surface leads to a temporary lowering of the moment of inertia and a slight increase in the rate of rotation. This glitch only lasts a few weeks until the star can resume its former structure. The crystalline nature of a neutron star surface is a new feature of these stars revealed by a study aimed at explaining the peculiar, sudden speed-up of some pulsars.

If we accept the fact, implied by the similarities between pulsars of widely different behavior, that all pulsars are similar objects and that the long period ones are simply older and have slowed down more, we must conclude that initially they form as neutron stars with densities $\gtrsim 10^{11}$ gm cm^{-3} and with radii $\lesssim$ 10 km rotating at periods as short as $\approx$ .020 sec, the extrapolated period of the Crab pulsar at its formation. Moreover, these neutron stars must possess enormously strong magnetic fields, of the order of 10^{10} gauss or more at the poles. This estimate follows from the reasonable assumption that magnetic flux is conserved as the star collapses in the supernova event that forms the neutron star. A star collapsing from the size of the Sun ($R_\odot \approx 7 \times 10^{10}$ cm) to the size of a neutron star ($R \approx 10^6$ cm) undergoes a change of radius by a factor 7×10^4 and a change of cross section area equal to the square of this factor, or $\approx 5 \times 10^9$. If the star had an initial polar field of the order of only 2 gauss, the neutron star field would increase to $\approx 10^{10}$ gauss. Many stars have much stronger initial fields, and if they are appreciably more massive than the Sun, as theory says they must be to become supernovae, then the final field may well exceed 10^{12} gauss. The presence of a strong field has several consequences of importance for the pulsar. It provides for polarization of the pulsar radiation in agreement with observation. It also provides a possible connection between the neutron star surface and a region some distance away from which the observed radiation may actually be emitted as required in some theories of pulsar emission. The field may also provide a sweeping beam or sheet of radiation to explain the reception of precisely timed pulses.

IV. Pulsar Emission Mechanisms

When we get down to details of possible mechanisms that might produce a beam or sheet of electromagnetic energy rotating with the star, it is no longer so simple to single out one of a number of competing theories. We shall have to be content to outline the basic requirements of a satisfactory theory and then to examine three different types of theories, briefly pointing out the advantages and weaknesses of each. It should be said at the outset that no fully satisfactory theory of pulsar electrodynamics has yet been developed, and it is not even certain that any of the three to be discussed here will prove in the long run to be correct, though each seems capable of explaining some pulsar features.

The energy flux emitted by a pulsar can be calculated in a straightforward fashion, and it gives an important clue to the kind of radiation mechanism to be sought. If a pulsar like the Crab emits on the order of 10^{38} ergs/sec from an object or radius $\approx$ 10 km, the flux emitted will be at least 10^{25} ergs cm^{-2} sec^{-1}, and probably more like 10^{26} erg cm^{-2} sec^{-1} since the pulsar doesn't radiate uniformly over its whole surface. A black body radiating this flux would have to have a temperature $\approx 2 \times 10^7$ °K (since Flux $= \sigma T^4$). This looks possible until one notes that the shape of the spectrum is definitely not that of a black body at 10^7 °K. If one considers the total energy in the radio spectrum ($\approx 6 \times 10^{30}$ ergs/sec for the Crab pulsar) and calculates the black body temperature required to produce this flux at radio frequencies, the resulting temperature is enormous, about 10^{20} °K. Only a coherent radiation source can possibly produce such radio fluxes. Thus one must seek a mechanism that will produce coherent oscillations like those of a tuned radio transmitter or a laser or maser. This coherence requirement puts a strong restriction on possible radiation mechanisms, but there appear to be at least two ways to satisfy this requirement, both of which depend upon the presence of a strong magnetic field.

One of these mechanisms is synchrotron radiation that results when relativistic electrons are injected into a strong magnetic field. The electrons follow helical paths about the magnetic lines of force and radiate at the frequency with which they spiral about the field lines. This is classical radiation in the sense that classical electromagnetic theory requires an accelerated charge to radiate. Naturally, the revolving charge suffers centripetal acceleration in its helical motion about a field line and radiates accordingly. The radiation is moderately coherent at the circular frequency of the electron. The latter slows down as the electron radiates away energy, leading to emission of a broad band of frequencies with the highest frequencies in the spectrum being proportional to the magnetic field strength. A field of 10^8 gauss will produce frequencies up into the far infrared while fields of 10^{12} gauss will give frequencies into the hard x-ray range ($\lambda \approx 1$ Å). Synchrotron radiation is polarized in the plane in which the electrons spiral, and relativistic effects cause it to be emitted anisotropically in a narrow-angle disk about the orbital plane. These features are shown in Figure 2 below. (See next page)

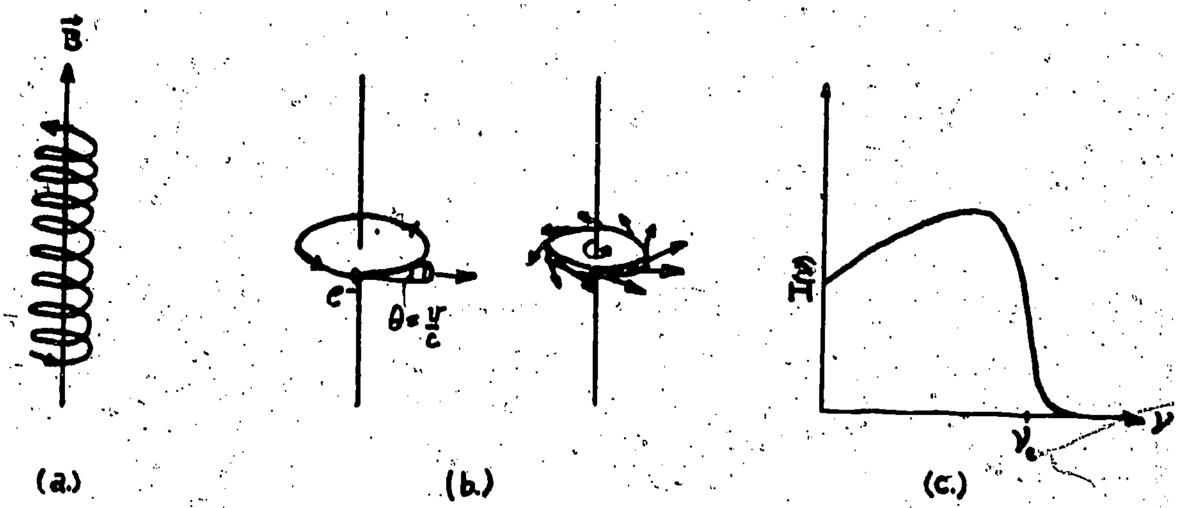


Figure 2. (a) Helical motion of an electron in a magnetic field.
(b) Synchrotron radiation of an orbiting electron.
(c) Spectral distribution of synchrotron radiation. ν_c is the cyclotron frequency, the orbiting frequency of the electron about the field lines.

The other type of coherent radiation to be considered involves a maser or laser type of emission but of a most unusual kind.(7, 8) Consider the radius of the circular or helical motion of an electron in a magnetic field. This radius is inversely proportional to the field strength, and if the field gets strong enough, the orbit radius can shrink to atomic size. When this happens, the orbit must obey quantum conditions just as it does in an atom. Once this quantum regime is reached, the electron no longer behaves classically, and in particular it no longer radiates synchrotron radiation. Now, as in the hydrogen atom, radiation can be emitted only when an electron jumps from one energy level to another. The levels can be calculated, as in the hydrogen atom, by requiring that a stable orbit have a circumference that is an integral number of deBroglie wavelengths of the electron. Detailed analysis along these lines shows that the energy gap between neighboring orbital states of an electron is 116 H/H_q KeV where $H_q = 4.4 \times 10^{13}$ gauss, the value of field at which the magnetic energy of the orbiting electron equals its rest energy (0.51 MeV). At low fields $<<H_q$ the energy levels of the electron in the field are virtually continuous since they are very closely spaced. But, at high enough fields, $H \approx H_q$, the levels are widely spaced, and the spacing may be large compared with kT. In this case, the excited levels are only sparsely occupied by thermally excited electrons. The system is much like an atom at low temperatures. If there is an efficient non-thermal means of populating the excited states of this system then the possibility for laser action exists, and this "magneto-laser" can in principle produce a strong, coherent beam of radiation as required. The necessary population inversion to cause laser action is postulated to be caused by electron collisions among energetic electrons accelerated by electrodynamic processes in the surface regions of the neutron star where there is a thin layer of "normal", non-degenerate matter. One such acceleration mechanism will be discussed below.

The rather exotic magnetic laser mechanism described above meets the requirements of pulsar emission in respect to brightness temperature, coherence, and anisotropy (beaming). The magnetic field also provides polarization. Electromagnetic radiation from such a laser is able to escape along the field lines, so it beams out from the magnetic poles. If the magnetic poles of the neutron star are not aligned with the axis of rotation, the radiation beam from the poles will sweep around in space like the beam from a lighthouse to give the characteristic pulse of the pulsar as seen by a distant observer. The figure below makes this clear. It also shows a case in which there are two pulses per revolution of the star (Figure 2b).

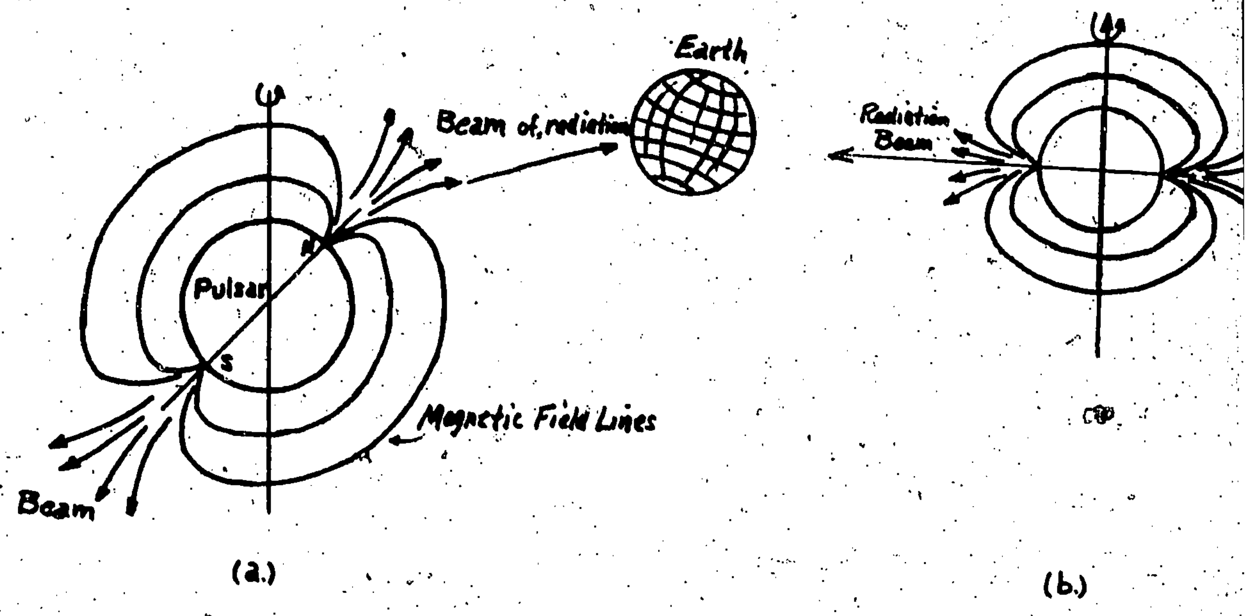


Figure 2. "Lighthouse" model of pulsar with (a) 1 pulse per revolution and (b) 2 pulses per revolution received on Earth.

Some pulsars do infact give "interpulses" which are weaker pulses between the main pulses. These may be a manifestation of Figure 2b.

One of the earliest theories of pulsar radiation was advanced by Gold[9] who proposed that the magnetic field lines rotate with the star as in Figure 2 above. Charged particles from the surface of the star are tied to the field lines, i.e., they rotate tightly about them. As these particles move away from the star along a given field line they are accelerated since the tangential velocity of a field line increases with distance from the axis of rotation according to the familiar relation $v_t = r\omega$. (See Figure 3 below). At a sufficient distance this tangential velocity approaches the speed of light and when this happens the particle radiates most of its accumulated kinetic energy as electromagnetic radiation. The entire process is a bit like the game of "crack-the-whip"; when the electron swings around fast enough, it lets go in a burst of electromagnetic radiation. The speed of light is attained on a cylinder of radius $r_c = c\omega^{-1}$ which is about 1,700 km for the Crab pulsar.

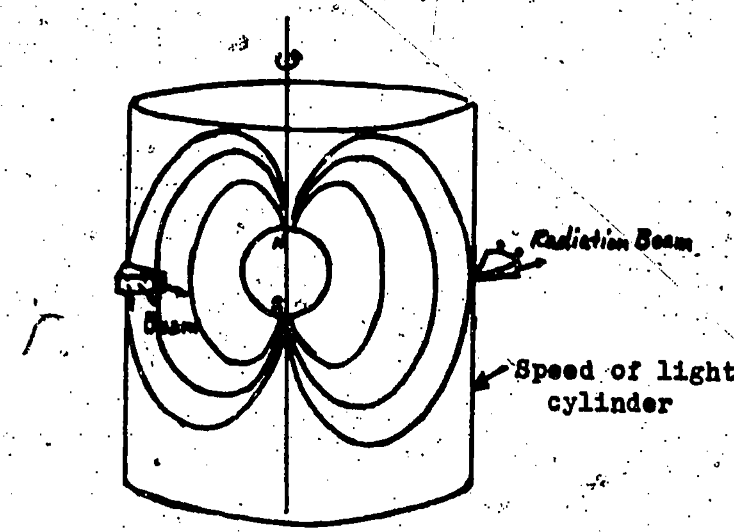


Figure 3. Gold's Model

The radiation at the speed of light cylinder will be somewhat beamed, as shown, by relativistic effects. The "beam" will actually be a continuous disk of radiation radiating in all directions from the equator of the cylinder as the situation is sketched in Figure 3. This would not produce pulses at a distant receiver. To do that, it must be postulated that electrons are emitted mainly either from localized "hot spots" on the surface of the neutron star or from its polar regions. These then follow the local field lines out to the speed of light cylinder where they radiate from a localized region on the cylinder. This region gives in turn a localized beam of radiation which races around the cylinder once each turn of the star. In this way the model can give pulses. The existence of a fairly strong field at the cylinder ($\approx 10^5$ gauss) provides polarization of the radiation. The polarization, frequency distribution, beam width, etc., are all characteristic of synchrotron radiation as would be expected since the electrons are being spun about in a giant circle by the revolving field lines and the electron orbits are similiar to those in a synchrotron.

One difficulty with the Gold model which is common to all models that propose a source of radiation located some distance away from the neutron star surface is that it is hard to see how the relation between source location and star surface can be maintained so accurately over a long period of time as to give the observed pulse rate precision. Any drift between the surface and the distant radiating region would spoil the precise repetition rate of the pulses.

Gold's model calls for relativistic electrons (≈ 3 MeV) to be delivered to the speed of light cylinder. This requires considerable acceleration of the electrons between the neutron star surface and the lightspeed cylinder. Two mechanisms that have been proposed to accomplish this acceleration will be described next.

Goldreich and Julian[10] suggested that the familiar mechanism of unipolar induction will operate to accelerate electrons. This can best be described in terms of the elementary laboratory demonstration of the Faraday disk dynamo. Consider a conducting disk rotating in a magnetic field as shown in Figure 4 below with slip rings on the axis and the edge of the disk to make electrical connections.

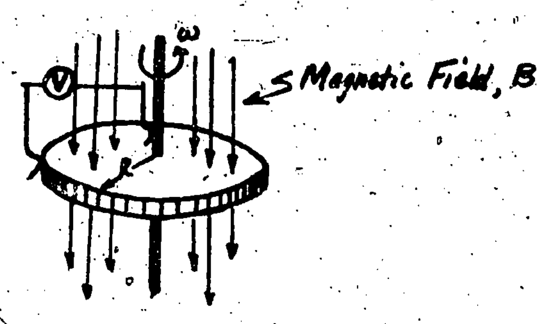


Figure 4. Simple Unipolar Generator

An electron of the rotating disk is carried across magnetic field lines with velocity $v = r\omega$ and experiences a radial force $F_r = evB = erB\omega$. This force is equivalent to a radial potential gradient $-e\frac{dV}{dr}$, or

$$- \int_0^V dV = \int_0^R B\omega\, r dr$$

$$V = \frac{1}{2} B\omega R^2 \qquad (8)$$

where V is the total voltage developed across the disk of radius R.

The important feature of this device for the pulsar application is that if electrons in the ionized gaseous plasma surrounding the rotating neutron star are not rigidly attached to the rotating magnetic field initially, then the field lines do pass by the electrons with some relative velocity and thereby induce an electric field, $-e\,\vec{v} \times \vec{B}$, that can accelerate the electrons. Detailed consideration of a dipole field shows that this mechanism is capable of accelerating electrons and ions away from the surface of a neutron star to energies well up into the cosmic ray range, far more than enough for the Gold model. The electric field generated in this manner is strong enough to separate the electrons and ions in the plasma near the star, and the charged particles move out along magnetic field lines as the diagram below shows.

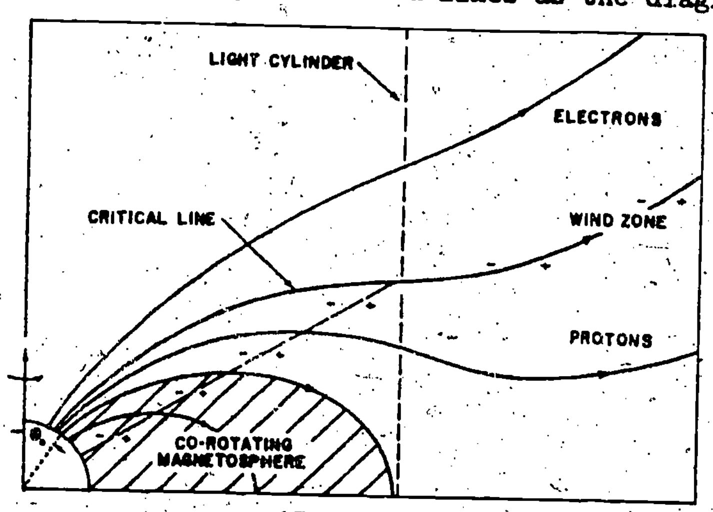


Figure 5. Schematic diagram of Goldreich-Julian model showing the corotating magnetosphere and stellar wind zone. The star is at the lower left.

The corotating magnetosphere is made up of charged particles attached to closed magnetic field lines. The field lines that cross the speed of light cylinder are open at the surface of the cylinder (they close farther out), and particles that radiate at the lightspeed surface follow these open field lines. The electrons and protons follow different sets of lines as a consequence of their difference in mass.

It can be seen from Figure 5 that there is a region on the light cylinder (where the particles intersect it) that comprises a source of radiation which sweeps around with the magnetic pole as the star rotates. This provides the required searchlight effect to produce observed pulses, and the radiation is of the synchrotron type at or near the light cylinder as in Gold's model. The Goldreich-Julian theory is an improvement, however, because it shows that a plasma and magnetosphere must exist near the surface of a neutron star instead of a vacuum, and the theory furnishes a method by which particles from this plasma may be accelerated away from the star to the lightspeed cylinder.

A different acceleration mechanism was proposed by Gunn and Ostriker(11) who pointed out that a rotating magnetic dipole would be a source of low frequency radiation that could accelerate electrons to high energies. A rotating magnetic dipole will generate in its vicinity a time-varying magnetic field. In itself such a field can accelerate electrons by the same method used in a betatron. Basic electromagnetic theory shows that a varying magnetic field generates an electric field $E \propto dB/dt$, and the field accelerates the particles. At distances beyond the lightspeed cylinder this time varying field becomes an electromagnetic wave for relativistic reasons. One might make the simplified argument that the field lines which carry energy cannot move faster than the speed of light, and when they reach that velocity they no longer co-rotate but move away at the speed of light as electromagnetic waves.

Classically the rotating dipole radiates at the rate

$$P = \frac{B_o^2 R^6 \omega^4}{12c^3} \text{ ergs/sec} \quad (9)$$

where B_o is the field strength at the pole of the star, R is the star radius, and ω is the rotational frequency. Taking $B_o \approx 10^{12}$ gauss, R = 10 km, and ω = 188 corresponding to a period of .033 sec, the total power radiated is $\approx 3 \times 10^{38}$ ergs/sec, approximately the total power onput of the Crab pulsar. Thus the rotating dipole model is capable of the required energy output. Charged particles are accelerated to the wave velocity near the star's surface and then ride the electromagnetic wave out to the light cylinder picking up additional energy from the wave in the process. The "wave riding" mechanism is similar to that in a linear accelerator and is the electromagnetic analog of surfing on an ocean wave.

There is a problem which plagues theories that postulate acceleration of particles near the star followed by transport to the light cylinder before radiation takes place. Classically, synchrotron radiation will take place in the acceleration region, and a calculation shows this process converts the energy gained by an electron into radiation so efficiently that the electron will not be able to reach relativistic speeds needed to reach the light cylinder by either of the mechanisms outlined above. At first glance this may seem to be of no importance: instead of waiting until it reaches the light cylinder, an electron simply radiates prematurely, and the net conversion of rotational energy into radiation is pretty much the same. However, in a dense emitting medium self absorption (i.e., reabsorption of emitted photons) takes place before the radiation can escape from the medium, and the resultant flux cannot exceed that of a blackbody at the temperature of the medium about one photon mean free path down into the medium from its "surface". Since the plasma surrounding a neutron star cannot exceed a temperature of the order of 10^9 °K, the escaping synchrotron radiation flux cannot exceed the blackbody flux from a source at that temperature. But the measured radiofrequency flux from a pulsar has a brightness temperature at least 10^{21} °K as pointed out earlier, hence the discrepancy introduced by self-absorption.

One way out of the self-absorption dilemma is offered by the magnetic laser mechanism of Chiu, Canuto, and Occhionero, mentioned earlier. Two factors act to eliminate the problem in this theory. When there is a population inversion with the upper energy levels of the system heavily populated, the system is no longer in thermodynamic equilibrium and the blackbody limit on radiated flux no longer applies. It is well known that the brightness temperature of laser emission is extremely high, or to put it another way, by emitting its radiation coherently a laser is able to exceed by many orders of magnitude the radiation output from a black body operating at the temperature of the lasing medium. The other factor mitigating the synchrotron self-absorption problem is that at the high fields under consideration the electron orbits about the field lines are quantized as mentioned earlier, and they do not radiate synchrotron radiation while in one of these quantized states. Moreover, if the field is high enough, the magnetic quantum states are far apart in energy, and only synchrotron radiation of very short wavelengths could be absorbed by such a quantized system. The longer wavelength radiation, if present, would travel out of the magnetosphere without absorption.

Only a few of many theoretical proposals for a mechanism of pulsar radiation have been discussed and these only very briefly. The subject is both complex and difficult, and to date no fully satisfactory theory has emerged.

lium from its "surface". Since the plasma
ot exceed a temperature of the order of 10^9°K,
ion flux cannot exceed the blackbody flux from
But the measured radiofrequency flux from a
ure at least 10^{21} °K as pointed out earlier,
l by self-absorption.

sorption dilemma is offered by the magnetic
, and Occhionero, mentioned earlier. Two
oblem in this theory. When there is a popu-
energy levels of the system heavily populated,
nodynamic equilibrium and the blackbody limit
es. It is well known that the brightness
extremely high, or to put it another way,
ntly a laser is able to exceed by many orders
t from a black body operating at the temperature
factor mitigating the synchrotron self-absorp-
fields under consideration the electron orbits
zed as mentioned earlier, and they do not
ile in one of these quantized states. More-
h, the magnetic quantum states are far apart
adiation of very short wavelengths could be
tem. The longer wavelength radiation, if
magnetosphere without absorption.

cal proposals for a mechanism of pulsar radia-
se only very briefly. The subject is both
te no fully satisfactory theory has emerged.

Those outlined above have various features in a successful theory, but each has its weak points as well. They serve to illustrate in some degree the wide variety of phenomena that must be considered in dealing with the phenomena and point out several directions one might expect a successful theory eventually to take.

References

(1) A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins, Nature, 217, 709 (1968).

(2) M. Ryle and J. A. Bailey, Nature, 217, 907 (1968).

(3) J. D. H. Pilkington, A. Hewish, S. J. Bell, T. W. Cole, Nature, 218, 126 (1968).

(4) C. R. Lynds, S. P. Maran, and D. E. Trumbo, Science, 161, 42 (1968).

(5) W. J. Cocke, M. J. Disney, D. J. Taylor, Nature, 221, 525 (1969).

(6) M. Ruderman, Nature, 223, 597 (1969).

(7) H. Y. Chiu, and V. Canuto, Phys. Rev. Letters, 22, 415 (1969).

(8) H. Y. Chiu, and F. Occhionero, Nature, 223, 1113 (1969).

(9) T. Gold, Nature, 218, 731 (1968).

(10) P. Goldreich, and W. H. Julian, Ap. J. 157, 869 (1969).

(11) J. E. Gunn, and J. P. Ostriker, Nature, 221, 454 (1969).